Not all synaesthetes are the same: Cognitive and personality differences in different types of synaesthetes

Anna Mas-Casadesús

Doctor of Philosophy in Psychology

School of Philosophy, Psychology, and Language Sciences
College of Arts, Humanities, and Social Sciences
The University of Edinburgh

October 2019

If [a man's] wit be no apt to distinguish or find differences, let him study the Schoolmen; for they are cymini sectore *[splitters of hairs].*

— Francis Bacon, *Of Studies*, 1597

Abstract

Synaesthesia is a perceptual condition in which stimulation of one sensory or cognitive pathway leads to automatic and involuntary experiences in a secondary sensory or cognitive pathway (e.g. seeing music or tasting words). Despite the fact that synaesthetes constantly perceive additional information during these inducer-concurrent associations, they are relatively unaffected by this irrelevant information. Chapter II investigates whether different samples of -visual synaesthetes (i.e. those experiencing synaesthesia types involving visual concurrents such as colours for letters or numbers – grapheme-colour synaesthesia – or sequence-space synaesthesias like calendar-forms) are better than-non synaesthetes at filtering out task-irrelevant stimuli in different conflict tasks. Synaesthetes were more efficient than controls at ignoring visual irrelevant stimuli presented together with tactile targets, but no group differences were observed when they had to perform the same visuo-tactile task with reversed instructions (i.e. attend visual and ignore tactile information) or in unimodal visual tasks (Studies 1 and 2). However, these results were not replicated in Study 3, which assessed a new sample of participants with the two versions of the same visuo-tactile tasks. This study also evaluated a) whether the observed synaesthetic attentional advantage was consistent across different sensory modalities combinations by introducing audio-visual modalities of the same tasks, and b) whether different types of -visual synaesthetes showed the same attentional advantages or not by comparing groups of colour-synaesthetes (i.e. those experiencing synaesthesias involving -colour as the concurrent) and sequence-synaesthetes (i.e. those experiencing sequence-space synaesthesias). Results revealed that sequence-synaesthetes were better than non-synaesthetes *and* colour-synaesthetes at filtering tactile irrelevant distractors presented with visual targets; no other group differences were observed. This suggests that the specific types of synaesthesias, together with other factors discussed, might play a relevant role in shaping the cognitive abilities of synaesthetes.

In order to explore the extent of the influence of synaesthetic individual differences, the second part on of the thesis examines differences in personality in individuals with different types of

synaesthesia (Chapter III – Study 4). Synaesthetes have a distinct personality profile compared to non-synaesthetes, but there are inconsistencies in the literature with respect to the personality traits that differ. Most studies have focused on grapheme-colour synaesthetes, ignoring other types of synaesthesia. Here, we compare matched groups of colour-synaesthetes, sequence-synaesthetes, and non-synaesthetes on the Big Five personality traits and specific empathy and positive schizotypy subscales. We replicated previous findings that synaesthetes experienced higher rates of Openness to Experience, Fantasising (a dimension of empathy), and Unusual Experiences (positive schizotypy) compared to non-synaesthetes. Importantly, some of these differences were only observed for sequence-synaesthetes, with higher rates of Openness to Experience compared to non-synaesthetes *and* colour-synaesthetes. However, no differences between synaesthetes and non-synaesthetes or between the two types of synaesthetes were found in a second sample assessed. We discuss several possible limitations affecting subject recruitment and assessment administration methods that could explain the different sample results.

The last section of the thesis addresses synaesthetic heterogeneity from a methodological point of view. The need to screen and classify synaesthetes led to the development and validation of a screening questionnaire, the Edinburgh Synaesthesia Screening Assessment or ESSA (Chapter IV – Study 5). Although synaesthetic tests of genuineness or consistency tests are considered the 'gold standard' of synaesthesia assessment, they are only available for a few synaesthesia types. The ESSA is a self-report questionnaire developed to cover an exhaustive range of synaesthesia types (108) and designed to assess both synaesthetes and non-synaesthetes by asking responders to rate how much each synaesthetic experience applies to them (5-point Likert scale). Sensitivity and specificity analyses were carried out on ESSA scores obtained from a sample of over 150 (synaesthete and non-synaesthete) participants who also completed synaesthetic consistency tests for -colour and sequence-space synaesthesias. Synaesthetes obtained significantly higher scores than non-synaesthetes, and the analyses showed acceptable rates of sensitivity and specificity (±85.5

and ±75.8, respectively). These results were internally and externally validated (in a new sample of 275 participants) yielding some modest values. We consider different detected bias and other factors that might reduce ESSA's performance and propose ways to address them in future studies.

In sum, converging evidence seems to indicate that synaesthetes are not a homogeneous category of individuals. Different cognitive and personality profiles are associated with different synaesthesia types. These findings have wider implications for the synaesthetic research area, as they suggest that grapheme-colour synaesthetes, predominantly assessed in in synaesthesia studies, might not be representative of all synaesthetes. These observations might at least in part explain contrasting results reported in the literature.

Lay Summary

Synaesthesia is a perceptual condition where a sensation in one of the senses, such as hearing, triggers a sensation in another sense, such as vision. For example, some people with synaesthesia (known as synaesthetes) see different colours when they hear or listen to music. Although synaesthetes constantly perceive these additional stimuli, they are relatively unaffected by it. The first part of this thesis investigates whether synaesthetes are generally better than people who do not have this condition or non-synaesthetes at filtering other non-synaesthetic irrelevant stimuli from the environment (e.g. distracting flashes of light). Specifically, we were interested in examining if synaesthetes would show advantages at ignoring irrelevant stimuli that matched the type of synaesthetic experiences that they have. In other words, if, for instance, synaesthetes who see colours for letters or numbers had an advantage at filtering (non-synaesthetic) visual irrelevant stimuli compared to non-synaesthetes.

To investigate this, our participants performed a series of tasks across different studies in which they had to respond to certain stimuli (targets) and ignore other (distractors). These target and distractor stimuli, always presented together, could be either visual (flashes of light of different colours), tactile (vibrotactile bursts on the fingers), or auditory (beep sounds). For example, in one of the tasks, participants were instructed to attend to green flashes of light and ignore vibrotactile bursts, or, in another, to attend to green flashes and ignore red ones. In a first study, we observed that synaesthetes appeared to be better than non-synaesthetes at ignoring visual distractors, but only when these were simultaneously presented with tactile targets. However, these results were not replicated in a subsequent study. Conversely, results revealed that synaesthetes were more efficient when they had to pay attention to visual targets while ignoring tactile distractors. In addition, only a specific type of synaesthetes, namely sequence-space synaesthetes or those individuals who experience synaesthesias such as number- or calendar-forms, seemed to show this filtering advantage in comparison to non-synaesthetes *and* other synaesthetes who had -colour experiences (e.g. colours for letters or

numbers). Therefore, this suggest that the specific types of synaesthesias experienced, together with other factors discussed, might play a relevant role in shaping the cognitive abilities of synaesthetes.

To explore the scope of the influence of synaesthetic individual differences, the second part of the thesis examines differences in personality traits between different types of synaesthetes. Previous studies have shown that synaesthetes have a distinct personality profile compared to non-synaesthetes, but there are inconsistencies in the literature with respect to the personality traits that differ. One possibility is that these differences are due to the presence of different types of synaesthetes. Here, we compare synaesthetes with -colour experiences, sequence-space synaesthetes, and non-synaesthetes on the Big Five personality traits and specific empathy and positive schizotypy subscales. We replicated previous findings that synaesthetes experienced higher rates of Openness to Experience, Fantasising (a dimension of empathy), and Unusual Experiences (positive schizotypy) compared to non-synaesthetes. Importantly, some of these differences were only observed for sequence-synaesthetes, with higher rates of Openness to Experience compared to non-synaesthetes *and* colour-synaesthetes. However, no differences between synaesthetes and non-synaesthetes or between the two types of synaesthetes were found in a second sample assessed. We discuss several possible limitations that could explain the different sample results.

The last section of the thesis addresses synaesthetic differences from a methodological point of view. The need to screen and classify participants for our studies led to the development and validation of a screening questionnaire, the Edinburgh Synaesthesia Screening Assessment (ESSA). Synaesthetic consistency tests, which measure how consistent is a person at reporting their specific synaesthetic associations (e.g. reporting that the letter 'A' is burgundy red over repeated times), are considered the 'gold standard' of synaesthesia assessment, but they are only available for a few synaesthesia types. The ESSA is a self-report questionnaire developed to cover an exhaustive range of synaesthesia types (108) and designed to assess both synaesthetes and non-synaesthetes by asking responders to rate how much each synaesthetic experience applies to them. We conducted a series of analyses aimed at assessing how sensitive and how specific the questionnaire was. That is, we evaluated how well the ESSA discriminated subjects with and without the condition. Over 150 (synaesthete and non-synaesthete) participants completed the questionnaire in addition to synaesthetic consistency tests for -colour and sequence-space synaesthesias. Synaesthetes obtained significantly higher scores than non-synaesthetes, and the analyses showed acceptable rates of sensitivity and specificity (±85.5 and ±75.8, respectively). These results were internally and externally (in a new sample of 275 participants) yielding some modest values. We consider different identified problems that might reduce ESSA's performance and propose ways to address them in future studies.

In sum, converging evidence seems to indicate that all synaesthetes are not the same. Different cognitive and personality profiles are associated with different synaesthesia types. These findings have wider implications for the synaesthetic research area, as they suggest that synaesthetes with colours for letters or numbers, predominantly assessed in in synaesthesia studies, might not be representative of all synaesthetes. These observations might at least in part explain contrasting results reported in previous studies.

Declaration

I declare that this thesis has been composed by myself and that the research reported here has been conducted by myself unless otherwise indicated. This work has not been submitted for any other degree or professional qualification. A modified version of Study 1 was published prior to the submission of this thesis (Mas-Casadesús & Gherri, 2017). The dissertation author was the primary investigator and author of this paper. In addition, a brief literature review on the topics covered in the Introduction section of Study 4 was published as part of the requirements for a research grant bursary received by the Psychology Postgraduate Affairs Group, UK (Mas-Casadesús, 2019).

- **Mas-Casadesús, A.** (2019). Is there a synaesthetic personality profile? Characterising synaesthetes' singularity. *PsyPAG Quarterly, 111*, 25-27.
- **Mas-Casadesús, A.**, & Gherri, E. (2017). Ignoring irrelevant information: Enhanced intermodal attention in synaesthetes. *Multisensory Research, 30*(3-5), 253-277.

Acknowledgements

As tempting as it could be to write the thank-you speech of a lifetime (let's be realistic, the Nobel prize is off the table), I will *try* to keep it brief.

First of all, I would like to thank my supervisor, Dr. Elena Gherri, for all her guidance, dedication, and support throughout these years. You told me once that I was very bad at receiving criticisms but very good at implementing the feedback. This sentence has stuck with me ever since. Because it defines me so well, and because I feel this has been the core of my relationship with the PhD – and with you. So thank you as well for keep (patiently) pushing me during all this time. I should also thank my second supervisor, Dr. David Carmel, whose passion for research is contagious; and Drs. Julia Simner and Jamie Ward of the University of Sussex for generously sharing with me part of the instrumental materials needed to conduct this research project.

I cannot thank enough my friends, for being there over the years and for listening to me go on about my research – and even more about all my ups and downs through this time. Eli, Guillem, Lena, Cristina, Laura, and Montse, thank you for making me forget about physical distances. Suzanne, for always giving me balance. Federica, Erminia, and Francesco, for embracing me into the Italian family. And, Ilaria, for being my constant constant. But, of course, none of this would have been possible without the support of my family in more ways I can enumerate. Mama, papa, Olga, sense el vostre suport, certament no estaria escrivint això ara mateix.

Lastly, I need to thank the School of Philosophy, Psychology, and Language Sciences and the College of Humanities and Social Sciences of The University of Edinburgh for economically supporting this research and this researcher with a 3-year PhD scholarship and numerous research support grants. I would also like to express my gratitude to the volunteer research assistant students who provided help in collecting data and to all the participants who took part in my studies.

Table of Contents

Abstract... i

Lay Summary.. v

Declaration...ix

Acknowledgments..xi

List of Tables xv

List of Figures xvii

1. Chapter I: General Introduction 1

1.1 What is synaesthesia? Basic concepts overview 1

1.2 Are synaesthetes different from non-synaesthetes? A selection of representative studies... 5

1.3 Are synaesthetes different from one another? Research questions and thesis outline 11

2. Chapter II: Attention and Multisensory Integration Processes in Synaesthesia 15

2.1 Chapter introduction 15

2.1.1 Do synaesthetes experience atypical multisensory integration?...................... 15

2.1.2 The role of attention and its complex interplay with multisensory integration... 20

2.1.3 Do synaesthetes experience atypical attention? 24

2.1.4 The cross-modal congruency task: A paradigm to study attention and multisensory integration processes in synaesthesia.. 28

2.2 Unimodal vs. cross-modal differences when vision is distractor (Study 1)................. 31

2.2.1 Introduction.. 31

2.2.2 Methods... 33

2.2.3 Results... 43

2.2.4 Discussion 48

2.3 Unimodal vs. cross-modal differences when vision is target (Study 2) 52

2.3.1 Introduction.. 52

2.3.2 Methods... 54

2.3.3 Results... 59

2.3.4 Discussion 62

2.4 Visuo-tactile vs. audio-visual differences: Vision as distractor vs. vision as target (Study 3).. 64

2.4.1 Introduction.. 64

2.4.2 Methods 66

2.4.3 Results 77

2.4.4 Discussion 83

2.5 Chapter discussion 89

3. Chapter III: Personality Trait Differences Between Synaesthetes (Study 4) 101

3.1 Introduction 101

3.2 Methods 105

3.3 Results 112

3.4 Discussion 116

4. Chapter IV: Development and Validation of the Edinburgh Synaesthesia Screening Assessment – ESSA (Study 5) 129

4.1 Introduction 129

4.2 Methods 134

4.3 Results 144

4.4 Discussion 155

5. Chapter V: General Discussion 163

References 173

Appendices 191

Appendix A: Edinburgh Synaesthesia Screening Assessment (ESSA Pilot Version) 193

Appendix B: Studies 1, 2, and 3 Extended Bayesian Interference Analyses 197

Appendix C: Studies 1, 2, and 3 Linear Mixed Model Analyses Additional Tables 201

Appendix D: Edinburgh Synaesthesia Screening Assessment (ESSA Revised Version) 209

Appendix E: Studies 1 and 3 Exploratory Analyses 219

Appendix F: Study 4 Exploratory Analyses 225

List of Tables

1. Descriptive, chi-square (χ^2), and *t*-statistics of Study 1 groups' demographics 36
2. Descriptive, chi-square (χ^2), and *t*-statistics of Study 2 groups' demographics 54
3. Descriptive, chi-square (χ^2), and *t*-statistics of Study 3 groups' demographics 70
4. -Colour and sequence-space synaesthetic consistency scores for Study 3 participants by group 70
5. Congruency effects values and statistics for each task of Study 3, as well as the visuo-tactile and audio-visual tasks combined, by group 79
6. Descriptive, chi-square (χ^2), and *t*-statistics of Study 4 groups' demographics 102
7. -Colour and sequence-space synaesthetic consistency scores of Study 4 participants, by sample and group 102
8. Mean scores by sample and group for the different personality subscales assessed in Study 4 106
9. Correlations between the personality subscales and the synaesthetic strength measures analysed in Study 4, by group (Sample A) 110
10. Descriptive, chi-square (χ^2), and *t*-statistics of Study 5 groups' demographics 131
11. Specifications regarding the different ESSA scorings calculated and analysed for the Brief and Extended modalities in Study 5 132
12. Specifications regarding the different ESSA direct statements analysed in Study 5 135
13. Mean scores by sample and group for the different ESSA scorings of Study 5 137
14. Mean scores by group for the different ESSA scorings of Study 5 (Training sample without other-synaesthetes) 138
15. Evaluation metrics for the different ESSA scorings of Study 5 (Training sample without other-synaesthetes) 138
16. Mean scores by group for the different ESSA scorings of Study 5 (Training sample with other-synaesthetes included in the synaesthete group).. 140
17. Evaluation metrics for the different ESSA scorings of Study 5 (Training sample with other-synaesthetes included in the synaesthete group).. 140
18. Mean scores by group for the different ESSA direct statements of Study 5 (Training sample).. 142
19. Evaluation metrics for the different ESSA direct statements of Study 5 (Training sample).. 142
20. Evaluation metrics for the 10-fold cross-validation analyses on the different ESSA scorings and direct statements of Study 5 (Training sample) 144
21. Mean scores by group for the different ESSA scorings and direct statements of Study 5 (Validation sample) 145

22. Evaluation metrics for the external validation analyses on the different ESSA scorings and direct statements of Study 5 (Validation sample) 146
C1. Linear mixed model analyses: model comparisons statistics (Study 1) 201
C2. Linear mixed model analyses: full model statistics (Study 1) 202
C3. Linear mixed model analyses: model comparisons statistics (Study 2) 203
C4. Linear mixed model analyses: full model statistics (Study 2) 204
C5. Linear mixed model analyses: model comparisons statistics (Study 3) 205
C6. Linear mixed model analyses: full model statistics (Study 3) 206
E1. Summary of the synaesthetic strength measures assessed in Studies 1 and 3 and their sources 222
E2. Correlations between the congruency effects and the synaesthetic strength measures analysed in Studies 1 and 3, by group 223

List of Figures

1. Study 1 Cross-modal Congruency Task design 38
2. Study 1 Flanker Task design 39
3. Study 1 Results 43
4. Study 2 Tasks design 56
5. Study 2 Results 59
6. Study 3 Audio-visual tasks design 73
7. Study 3 Results – Mixed Analyses of Variance 76
8. Study 3 Results – Linear mixed model analyses 81
9. Study 3 Exploratory analyses results – Both-synaesthetes 85
10. Study 3 Exploratory analyses results – Other-synaesthetes 87
11. Study 1-3 Meta-analysis output xxx
12. Study 1-3 Meta-analysis influence analysis output xxx
13. Study 4 Results – Group differences (Sample A) 108
14. Study 4 Results – Synaesthetic strength (Sample A) 110
15. Study 4 Exploratory analyses results – Both-synaesthetes (Sample A) 114
16. Study 4 Exploratory analyses results – Other-synaesthetes (Sample A) 115
17. Study 4 Exploratory analyses results – Other-synaesthetes (Sample B) 116
18. Study 5 Results – ESSA responses by synaesthetic categories 137
19. Study 5 Results – ESSA scorings ROC curves (Training sample without other-synaesthetes) 139
20. Study 5 Results – ESSA scorings optimal thresholds (Training sample without other-synaesthetes) 139
21. Study 5 Results – ESSA scorings ROC curves (Training sample with other-synaesthetes) 141
22. Study 5 Results – ESSA scorings optimal thresholds (Training sample with other-synaesthetes) 141
23. Study 5 Results – ESSA direct statements ROC curves (Training sample) 143
24. Study 5 Results – ESSA direct statements optimal thresholds (Training sample) 143
25. Study 5 Results – ESSA scorings and direct statements optimal thresholds (Validation sample) 146

1. Chapter I: General Introduction

This general introduction provides an overview on the basic concepts and scientific literature on synaesthesia and serves as a basis for the specific theoretical contents that will be covered in this thesis, with more detailed reviews in the chapters that follow.

1.1 What is synaesthesia? Basics concepts overview

Synaesthesia, from the Ancient Greek *syn* ("together") and *aisthēsis* ("sensation"), is a rare, non-pathological phenomenon in which the experience of an attribute of a stimulus, known as the inducer (e.g. its shape, sound, or meaning), involuntarily induces the conscious perception of another attribute within the same or different modality, called the concurrent (e.g. Ward, 2013). For instance, in grapheme-colour synaesthesia, seeing the letter "A" in black ink automatically triggers a red photism (e.g. Simner, 2012). Synaesthetes can experience this colour superimposed on the ink (i.e. outside their body; 'projector' synaesthesia) or in their mind's eye (i.e. 'associator' synaesthesia) (Dixon, Smilek, & Merikle, 2004). Some synaesthesia types are only activated by highly specific perceptual inducers (also called 'low' synaesthesia; e.g. Arnold, Wegener, Brown, & Mattingley, 2012; Jansari, Spiller, & Redfern, 2006; Witthoft & Winawer, 2006). For instance, in sound-colour synaesthesia, hearing particular sounds or music causes the perception of different colour experiences. However, most synaesthesia types are triggered by the attribute's general concept regardless of the sensory modality that activates it (i.e. 'high' synaesthesias; e.g. Chiou & Rich, 2014; Rich, Bradshaw, & Mattingley, 2005). That is, either seeing, hearing, or thinking about the inducer can trigger the concurrent. For example, in sequence-space synaesthesia, ordinal sequences such as the days of the week, months, or numbers are automatically seen as visuo-spatial representations arranged in specific forms like circles or lines (e.g. Jonas & Price, 2014). These visualisations are activated both when, for instance, these synaesthetes think about their doctor's appointment the following Tuesday or the shop assistant tells them how much they need to pay.

Over 70 types of synaesthesias have been documented to date (Day, 2019). Moreover, synaesthetic inducer-concurrent associations are not only highly arbitrary and specific, but they are also largely idiosyncratic. This means that, for example, two grapheme-colour synaesthetes might not share any colours for all the given letters of the alphabet. Or that distinctions such as 'crimson red' or 'burgundy red' are of relevance to the individual. However, all synaesthesia types share a number of characteristics that have become to be known as the synaesthetic defining criteria (see Simner, 2012 and Ward, 2013 for reviews). Notably, as mentioned above, synaesthetic associations occur automatically and involuntarily: i.e. synaesthetes do not try to experience their synaesthesias but, on the contrary, they cannot avoid experiencing the concurrents once they perceive the inducers. Secondly, despite the high specificity of synaesthetic inducer-concurrent pairings, these tend to remain stable or constant for the subject over time. The synaesthete who described the letter "A" as 'crimson red' will report the same exact shade of red for this letter if asked again at a later date. In fact, most synaesthetes tend to report experiencing their synaesthesia for as long as they can remember and records of consistency over long periods of time up to years are documented (e.g. Dresslar, 1903; Baron-Cohen, Burt, Laittan-Smith, Harrison, & Bolton, 1996; Baron-Cohen, Harrison, Goldstein, & Wyke, 1993; Baron-Cohen, Wyke, & Binnie, 1987; Ginsberg, 1923; Simner & Logie, 2008 – but see e.g. Meier, Rothen, & Walter, 2014; Simner, Harrold, Creed, Monro, & Foulkes, 2009a; Simner, Ipser, Smees, & Álvarez, 2017 for recent evidence questioning this stability across the adult lifespan).

Internal synaesthetic consistency is considered such a robust feature that it has become the 'gold-standard' of synaesthetic assessment. The so-called tests of genuineness or consistency tests measure how consistent are people at reporting their specific inducer-concurrent associations. For example, grapheme-colour synaesthetes are asked to report their colours for each letter of the alphabet or sequence-space synaesthetes to draw or describe their number lines. This process is repeated several times in the same session or in different sessions after determined periods of time (sometimes up to years) and answers are then compared. Synaesthetes, compared to non-synaesthetes, tend to show very low deviance or great consistency in their different answers (e.g. 92.3% consistency over a year period for synaesthetes in contrast to a 37.6% for controls tested only after a week; Baron-Cohen et al., 1993). Initially, these assessments took the form of simple test-retests and, over the years, the methodological approaches and materials have been gradually standardised, proving to be powerful and reliable tools to distinguish between synaesthetes and non-synaesthetes. However, there are some concerns related to the stability of consistency itself (see above) and, most importantly, only a few types of synaesthesias can be currently assessed through consistency tests (i.e. grapheme-colour, music-colour, sequence-space, ordinal-linguistic personifications, lexical-gustatory, and – saving some methodological distances – mirror-touch synaesthesias; see section 4.1). For that reason, ad-hoc self-report questionnaires and interviews are often used as complimentary approaches. The utility (and necessity) of self-report is widespread across diverse fields of empirical research, but these types of data are subjected to intrinsic biases or limitations related to reliability which must also be considered (Chan, 2009).

The different assessment and screening methods described have been used to study other aspects of synaesthesia such as population prevalence estimates. The condition is suggested to be present in (at least) about 4.4% of the population, equally experienced by both sexes (Simner et al., 2006). However, this has been a topic of on-going controversy in the field. Investigations addressing this area of study have offered prevalence estimates ranging from 0.004% (Cytowic, 1989) to 26% (Mann, Korzenko, Carriere, & Dixon, 2009). Similarly, female-to-male ratios have oscillated from the currently accepted 1:1 to traditionally predominantly female projections with ratios up to 6.2:1 (Rich et al., 2005) (see Johnson, Allison, & Baron-Cohen, 2013 for a summary of all these studies). One of the main factors causing these big disparities is the fact that there is not a clear definition of what is synaesthesia (and what is not). This ambiguity influences which types are included in prevalence screenings and, consequently, final incidence estimates. Secondly, almost all prevalence studies have recruited participants via self-referral. This does not only leave out people who do not make an effort to come to the lab, but also those who do have been suggested to have particular characteristics and might therefore not be entirely representative of the synaesthetic population at large (Carmichael, Down, Shillcock, Eagleman, & Simner, 2015). In addition, women are known to be more prone to self-referral than men, explaining as well issues related to gender biases (Simner et al., 2006). Lastly, other sampling biases, methodological differences across studies (e.g. use of consistency tests or not, screening of specific synaesthesia types or not), or the composition of the cohorts studied are further factors to consider.

There are several documented instances of synaesthesia acquired due to external causes such as traumatic head injury, stroke, or loss of sensory input, among others (e.g. Afra, Funke, & Matsuo, 2009; Armel & Ramachandran, 1999; Cacace et al., 1999; Fornazzari, Fischer, Ringer, & Schweizer, 2012; Goller, Nowak, Richard, & Ward, 2011; McFeely, Antonelli, Rodríguez, & Holmes, 1998; Ramachandran, Rogers-Ramachandran, & Cobb, 1995; Ro et al., 2007; Schweizer et al., 2013). Synaesthesia has also been induced (or attempted to) through drugs, hypnosis, and training (see Terhune, Luke, & Cohen-Kadosh, 2017 for a review on the topic). However, most cases, and the ones mainly reflected in prevalence estimates and in the scientific literature in general, are of developmental synaesthesia. Developmental synaesthesia is considered to be present from birth and not as a result of neurological damage or any other extraneous events. Members of a same family tend to show the condition, but not necessarily the same synaesthetic types or the same specific inducer-concurrent associations (e.g. Barnett et al., 2008; Jewanski, Simner, Day, & Ward, 2011). In addition, it has been observed that having one type of synaesthesia is linked to an increased probability of experiencing additional types, especially within the same family of concurrents (e.g. Cytowic & Eagleman, 2009; Niccolai, Jennes, Stoerig, & van Leeuwen, 2012; Novich, Cheng, & Eagleman, 2011; Sagiv, Simner, Collins, Butterworth, & Ward, 2006b). The study of synaesthesia's genetic heritability and development is still in its infancy, but recent converging evidence points to a complex genetic heterogeneity with multiple gene candidates and different modes of inheritance that would potentially explain the different phenotypes observed within families and in the population (see Asher & Carmichael, 2013 for an in-depth review).

Understanding which genes cause or promote synaesthesia will also help elucidating the neurological mechanisms underlying the condition. Different models have been proposed. The cross-activation model (Ramachandran & Hubbard, 2001a) states that synaesthesia is the result of direct increased connectivity between the cortical sensory areas involved in the specific inducer-concurrent associations. According to this theory, synaesthetes' brains would be thus structurally different from non-synaesthetes. On the other hand, the disinhibition feedback model (Cohen-Kadosh & Walsh, 2008; Grossenbacher & Lovelace, 2001) and the re-entrant feedback model (Smilek, Dixon, Cudahy, & Merikle, 2001) suggest that synaesthesia occurs due to a lack of inhibition of the feedback from higher cortical areas (in particular, the parietal lobe) failing to suppress irrelevant activation from the synaesthetically implicated lower cortical areas. Hence, in this case, synaesthetes' brains would show functional differences compared to non-synaesthetes (Cohen-Kadosh, Henik, Catena, Walsh, & Fuentes, 2009). Acknowledging the developmental nature of synaesthesia and considering the genetic evidence thus far, Carmichael and Simner (2013) additionally proposed that early life interactions between the central nervous system and the immune system, through genes involved with both brain development and immunity, might play a role on how these peculiar cortical connections arise. Although these different frameworks present conceptual differences, it is important to notice that they are not mutually exclusive. Moreover, they all involve differences in connectivity compared to the neurotypical brain.

1.2 Are synaesthetes different from non-synaesthetes? A selection of representative studies

A number of investigations (mostly conducted on grapheme-colour synaesthetes) have observed structural and functional brain differences between synaesthetes and non-synaesthetes. Specifically, synaesthetes appear to have greater volumes of grey and white matter in the fusiform gyrus in the V4 colour area, the neighbourhood area of V4, and in the primary auditory cortex as well as other concurrent brain areas (Hänggi, beeli, Oechslin, & Jäncke, 2008; Jäncke, Beeli, Eulig, & Hänggi, 2009; Weiss & Fink, 2008). These studies also observed larger superior parietal lobes for synaesthetes than controls. Lastly, Hänggi et al. (2008) and Jäncke et al. (2009) found an increase in grey matter density of sensory areas that were not necessarily related to the sensory modalities of synaesthetic experiences (i.e. increased grey matters in V1 and V2 and in the secondary somatosensory cortex). These widespread differences appear to be evident as well in terms of functional brain differences. Bilateral activation in the occipito-temporal cortex, including, but not restricted to, the V4 colour area, has been observed in several studies (Laeng, Hugdahl, & Specht, 2011; Nunn et al., 2002; Rouw & Scholte, 2007; Steven, Hansen, & Blakemore, 2006; Weiss, Zilles, & Fink, 2005). Importantly, whole-brain analyses also report increased clusters of activation of the parietal cortex and near the intraparietal sulcus and in the angular gyrus (Laeng et al., 2011; Nunn et al., 2002; Paulesu et al., 1995; Rouw & Scholte, 2010; Steven et al., 2006; Weiss et al., 2005). In addition, several studies have observed increased activation of several areas during the synaesthetic experience: the bilateral insula and operculum (Nunn et al., 2002; Paulesu et al, 1995; Sperling et al., 2006), the left precentral gyrus (Laeng et al., 2011; Paulesu et al., 1995; Nunn et al., 2002; Rouw & Scholte, 2010; Weiss et al., 2005), and the frontal lobe (Laeng et al., 2011; Paulesu et al., 1995; Sperling, Prvulovic, Linden, Singer, & Stirn, 2006) (see Rouw, Scholte, & Colizoli, 2011 for information on the specific studies and a detailed review on the topic, and Dojat, Pizzagalli, & Hupé, 2018 and Hupé & Dojat, 2015 for counter critiques).

Genetic and brain structural/functional differences between synaesthetes and non-synaesthetes are also reflected at a cognitive level. A great number of investigations have shown differences in cognition between synaesthetes and non-synaesthetes. For instance, there is evidence that grapheme-colour synaesthetes have memory advantages on tests of colour recognition memory as well as tests that involve recalling both inducer- and concurrent-related materials (Gross, Neargarder, Caldwell-Harris, & Cronin-Golomb, 2011; Lunke & Meier, 2018; Meier & Rothen, 2007; Radvansky, Gibson, & McNerney, 2011; Rothen & Meier, 2010; Rothen, Nyffeler, von Wartburg, Müri, & Meier, 2010; Rothen, Meier, & Ward, 2012; Teichmann, Nieuwenstein, & Rich, 2017; Ward, Hovard, Jones, & Rothen, 2013; Yaro & Ward, 2007). In addition, general memory benefits for materials unrelated to synaesthesia have also been reported (Bankieris & Aslin, 2016a; Gross et al., 2011; Lunke & Meier, 2018; Pritchard, Rothen, Coolbear, & Ward, 2013; Rothen & Meier, 2010; Rothen, Meier, & Ward, 2012; Ward et al., 2013). Only very few studies have examined memory abilities in other types of synaesthetes, but similar advantages have been observed for sequence-space synaesthetes (Brang, Teuscher, Ramachandran, & Coulson, 2010; Lunke & Meier, 2018; Simner, Mayo, & Spiller, 2009b) and sound-colour synaesthetes (Lunke & Meier, 2018). Lunke and Meier (2018) is also the first study known to date that has assessed and compared different types of synaesthetes (grapheme-colour, sound-colour, grapheme-and-sound-colour, and sequence-space). They observed a consistent general memory advantage (i.e. synaesthetetic unrelated material) for all the types of synaesthetes evaluated, but only grapheme-colour showed concurrent-specific benefits and only grapheme-and-sound-colour showed inducer-specific benefits, suggesting thus different mechanisms for different types of synaesthetes.

These synaesthetic inducer and concurrent memory advantages would be related to evidence from other studies that have suggested enhanced perceptual processing for synaesthetes associated with the domain of their synaesthesia. For example, several studies have shown superior colour processing in synaesthetes who experience colour concurrents compared to non-synaesthetes (Arnold et al., 2012; Banissy et al., 2013b; Banissy, Walsh, & Ward, 2009; Yaro & Ward, 2007). At the same time, other reported differences in early visual processing with tasks not directly related to synaesthesia could explain the observed general memory benefits. Grapheme-colour synaesthetes have been found to have increased visual-evoked potentials to high-spatial frequency Gabor patches (selectively biasing parvocellular pathway responses, related to colour processing; Barnett et al., 2008); lower phosphene thresholds in response to occipital lobe stimulation (i.e. visual cortex hyperexcitability; Terhune, Tai, Cowey, Popescu, & Cohen-Kadosh, 2011); and worse motion coherence processing abilities compared to non-synaesthetes (interpreted as a bias caused due to favouring colour processing over motion perception in a context of resource competition; Banissy et al., 2013b). Lastly, Ward, Rothen, Chang, and Kanai (2017a) assessed a broad range of visual abilities comparing, for the first time, different types of synaesthetes (grapheme-colour, sequence-space, and grapheme-colour-and-sequence-space). Amongst other findings, they replicated the observation that all these groups of synaesthetes showed increased abilities to discriminate colour. The authors also established the novel finding that all of them showed enhanced shape/curvature perception, which has been observed to be an important feature in primate V4 neurons (important for colour perception). But perhaps the most relevant finding for the aims of this thesis was that only those synaesthetes who only experienced sequence-space synaesthesia showed an advantage in high-spatial frequency perception.

Several authors argue that mental imagery is mediated, at least in part, by the same neural circuits in the brain that process perceptual information coming from the outside word (e.g. Borst & Kosslyn, 2008; Kosslyn, 1994; Thompson, Slotnick, Burrage, & Kosslyn, 2009). Interestingly, Francis Galton, one of the first authors to scientifically document synaesthesia, considered the condition just a heightened expression of mental imagery (Galton, 1880). Nowadays, synaesthesia is regarded as a distinct phenomenon, but considering that a great number of synaesthetic experiences can be internally triggered (i.e. ‘high’ synaesthesia or synaesthesia elicited by the attribute’s general concept), it might share some mechanisms with standard mental imagery (O’Dowd, Conney, McGovern, & Newell, 2019; Price, 2013; Spiller, Harkry, McCullagh, Thoma, & Jonas, 2019). A number of studies have indeed observed that synaesthetes in general, and sequence-space synaesthetes in particular, seem to experience higher rates of self-reported visual imagery (Barnett & Newell, 2008; Chun &

Hupé, 2016; Havlik, Carmichael, & SImner, 2015; Janik McErlean & Banissy, 2016; Price, 2009; Rizza & Price, 2012; Simner et al., 2009b; Spiller & Jansari, 2008; Spiller, Jonas, Simner, & Jansari, 2015). However, there is contradictory evidence supporting the association of these self-reports with observations of better than average visuo-spatial skills in behavioural tasks (Havlik et al., 2015; Rizza & Prize, 2012; Simner et al., 2009b; Spiller & Jansari, 2008). Some authors have suggested that part of these discrepancies could be due to individual differences within synaesthetes such as synaesthetic strength (particularly, the number of synaesthesia types experienced; Havlik et al, 2015; Spiller et al., 2015).

All these differences between synaesthetes and non-synaesthetes are framed into a wider context of evidence which includes other research areas like multisensory integration or attention processes (which will be dealt with in the next Chapter) that seems to reinforce the idea of a general ‘synaesthetic perceptual and cognitive style’. But synaesthetes’ singularity appears to extend to other domains. For instance, experiencing the world through synaesthesia might cause specific personality characteristics or, alternatively, having synaesthesia might predispose these individuals to certain personality traits in the first place through common genetic factors (Banissy et al, 2013a). Evidence seems to support this hypothesis. Previous studies have shown that synaesthetes seem to have an atypical personality profile compared to non-synaesthetes consistently characterised by higher rates of Openness to Experience (Banissy et al., 2013a; Chun & Hupé, 2016; Rouw & Scholte, 2016); of Fantasising, a dimension of empathy (Banissy et al., 2013a; Chun & Hupé, 2016; Rader & Tellegen, 1987; Rouw & Scholte, 2016); of and Unusual Experiences (positive schizotypy; Banissy et al., 2012; Janik McErlean & Banissy, 2016). Synaesthesia has also been less consistently linked to higher rates of Neuroticism (Banissy et al., 2013a; Rouw & Scholte, 2016) and Emotionality (Banissy & Ward, 2007; Rouw & Scholte, 2016), and lower rates of Agreeableness (Banissy et al., 2013a) and Conscientiousness (Rouw & Scholte, 2016). However, divergences to this general synaesthetic personality profile have been observed for music-colour and sequence-space synaesthetes (Rader & Tellegen, 1987 and

Ward et al., 2018a, respectively). In addition, some studies have suggested the possibility that synaesthetes who report stronger synaesthetic experiences (as measured with consistency tests) or report a greater number of synaesthesia types, as opposed to those who present weaker experiences or fewer types, might also show differences with respect to the synaesthetic personality profile. Rouw and Scholte (2016) and Hossain, Simner, and Ipser (2018) investigated this and provided initial evidence indicating that there might be a positive relationship between the strength or number of synaesthesias and the intensity of the personality traits experienced.

Synaesthesia might also have shared genetic or neurological basis with several developmental, physical, and mental health disorders. There is evidence showing links between synaesthesia and autism, with higher prevalence of (grapheme-colour) synaesthesia in autistic populations (Baron-Cohen et al., 2013; Neufeld et al., 2013) compared to the general population (Simner & Carmichael, 2015). However, some researchers observed that these co-occurring rates only happened for those people within the autism spectrum disorder who also have savant skills (Hughes, Simner, Baron-Cohen, Treffert, & Ward, 2017). In addition, stronger synaesthesia (i.e. higher synaesthetic consistency scores) seems to be associated with higher rates of autistic traits (Burghoorn, Dingemanse, van Lier, & van Leeuwen, 2019). On another note, Carmichael, Smees, Shillcock, and Simner (2019) recently conducted a large-scale study, the most ambitious of its kind to date, screening almost 4,000 people on grapheme-colour synaesthesia and with a health questionnaire which included 24 conditions representative of different population clinical disorders. The authors found that grapheme-colour synaesthesia was comorbid with anxiety disorder. Importantly, they did not replicate previously suggested links between synaesthesia and irritable bowel syndrome (Carruthers, Miller, Tarrier, & Whorwell, 2012), migraine (Jonas & Hibbard, 2015; Jürgens, Schulte, & May, 2014), multiple sclerosis (Simner, Carmichael, Hubbard, Morris, & Lawrie, 2015), or autism spectrum disorder (Baron-Cohen et al., 2013; Neufeld et al., 2013; but refer back to Hughes et al., 2017 for a possible explanation).

1.3 Are synaesthetes different from one another? Research questions and thesis outline

The studies reviewed so far have shown that there are clear differences between synaesthetes and non-synaesthetes which are present with different degrees of intensity and at different levels of explanation: from genes, to brain structure/function, or to perception and cognition, in addition to personality and other domains of affectation such as physical or mental illness. While these studies have typically considered synaesthetes as a homogenous group, the basis of the synaesthetic phenomenology is intrinsically heterogenous. For instance, we can differentiate between associator and projector or high vs. low synaesthetes. Behavioural evidence seems to support this diversity as well, suggesting that for low/associator synaesthetes, compared to high/projector synaesthetes, the early effects of synaesthesia might occur completely prior to awareness (see van Leeuwen, 2013 for a review on the topic). Further supporting this hypothesis, a neuroimaging study found that projector grapheme-colour synaesthetes tended to show bottom-up modulation (i.e. attentional guidance driven by external rather than internal factors) when presented by synaesthesia-inducing graphemes (van Leeuwen, den Ouden, & Hagoort, 2011). However, what level of awareness is necessary to elicit synaesthesia, or what is the general role of attention in synaesthesia (see Rich & Mattingley, 2013), are still topics of on-going debate in the field.

In addition, even though there are over 70 currently reported types of synaesthesia (Day, 2019), the vast majority of research on synaesthesia has been conducted on grapheme-colour synaesthesia, as it is one of the most prevalent types (Simner et al., 2006). Moreover, the presence or absence of additional synaesthesia types and the possible impact that this could have on the variables of interest studied, has been largely ignored. However, there are some initial indications that these factors might be of importance. As seen in the previous section, only a few studies have directly assessed and compared different types of synaesthetes, but they have importantly shown specific characteristics for particular types. For instance, in the memory study of Lunke and Meier (2018), different subgroups of synaesthetes were observed to have differences with respect to specific memory benefits related to synaesthetic stimuli. Or Ward et al. (2017a) found that only sequence-space synaesthetes showed advantages at high-spatial frequency perception compared to grapheme-colour and grapheme-colour and sequence-space synaesthetes. It is worth noting that better performance for this type of visual ability had been previously detected in a (supposed) group of grapheme-colour synaesthetes (Barnett et al., 2008). But given Ward et al.'s (2017a) findings, it is likely that (at least) some of the individuals of Barnett et al.'s (2008) sample experienced sequence-space synaesthesia as well (this information was not reported). Sequence-synaesthetes have also been observed to experience higher rates of self-reported visual imagery compared to other synaesthetes (see previous section for references) and to show differences to the general synaesthetic personality profile (Ward et al., 2018a). Lastly, other individual differences within synaesthetes such as synaesthetic strength also seem to be relevant. Synaesthetic strength, measured as the number of synaesthesia types reported or the degree of synaesthetic consistency experienced (i.e. consistency scores), might modulate the rates of self-reported visual imagery (Havlik et al., 2015; Spiller et al., 2015) and personality traits (Hossain et al., 2018; Rouw & Scholte, 2016), or the relationship between synaesthesia and autistic traits (Burghoorn et al., 2019).

It is possible then that poor replication rates in the field of synaesthesia research might be in part due to sampling issues whereby all synaesthetes were treated and considered as a uniform category of individuals. Therefore, the primary aim of this thesis is to contribute to the study of differences between different groups of synaesthetes. Moreover, in an attempt to explore the scope of the implications of these individual differences, this will be done from various points of view and areas of research. Chapter II addresses differences in attention and multisensory integration processes in synaesthesia. Both cross-modal correspondences and the role attention in synaesthesia have attracted a great deal of research, but, to date, only a few studies have considered how these two cognitive processes might relate to each other in synaesthesia. In particular, Chapter II investigates differences at filtering out task-irrelevant stimuli in different cross-modal and unimodal conflict tasks comparing -visual synaesthetes (i.e. those synaesthetes experiencing at least one synaesthesia type involving visual concurrents; e.g. grapheme-colour or sequence-space synaesthetes) to non-synaesthetes in Studies 1 and 2 and within synaesthetes (colour- vs. sequence-space synaesthetes) and non-synaesthetes in Study 3.

In order to explore the extent of the influence of synaesthetic individual differences, Chapter III (Study 4) examines differences in personality traits between different types of synaesthetes. Previous evidence has shown that synaesthetes have a distinct personality profile compared to non-synaesthetes, but there are inconsistencies in the literature with respect to the personality traits that differ. One possibility is that these differences are due to the presence of different types of synaesthetes. Here, we compare synaesthetes with -colour experiences (i.e. those synaesthetes experiencing at least one synaesthesia type involving colour concurrents), sequence-space synaesthetes, and non-synaesthetes on the Big Five personality traits and on specific empathy and positive schizotypy subscales. The possible effects of other types of synaesthesias and other individual difference factors such as synaesthetic strength are also evaluated throughout the different topics covered in Chapter II and III.

Finally, Chapter IV (Study 5) addresses synaesthetic heterogeneity from a methodological point of view. The predominance of grapheme-colour in synaesthesia research has also led to limitations and biases in terms of synaesthetic screening and assessment. Grapheme-colour synaesthesia (and some other -colour experiences) can be assessed with synaesthetic tests of genuineness or consistency tests, considered the 'gold standard' of synaesthesia assessment, but similar tests are only available for a few other synaesthesia types. This means that the majority of synaesthesias cannot be objectively measured and thus individuals with less frequent types of synaesthesias can be potentially misclassified, with the confounding risks that this implies. Self-report interviews and questionnaires can be good methodological tools to screen both synaesthetes and non-synaesthetes and address synaesthetic variability at the same time, but no standardised such measurement exists to date. For that reason, Study 5 presents the development and validation of a new self-report synaesthesia screening questionnaire, the Edinburgh Synaesthesia Screening Assessment (ESSA).

2. Chapter II: Attention and Multisensory Integration Processes in Synaesthesia

2.1 Chapter introduction

2.1.1 Do synaesthetes experience atypical multisensory integration?

One of the main questions with respect to synaesthesia is whether it constitutes a distinct phenomenon that can be qualitatively distinguished from typical perception or it is an enhanced experience of normal multisensory integration (e.g. see Newell & Mitchell, 2016 for a detailed review). Even though synaesthetic experiences are highly idiosyncratic (i.e. two grapheme-colour synaesthetes might have two completely different colours for the same letter), they often reflect patterns of typical cross-modal associations observed in the general population (Bankieris & Simner, 2015). For instance, both synaesthetes and non-synaesthetes tend to associate bright colours or lightness to high pitches and dark colours to low pitches (e.g. Marks, 1974; Ward, Huckstep, & Tsakanikos, 2006), and the two populations have similar preferences for certain pairings between colours and letters or colours and days of the week (e.g. red for 'A' and 'Monday' and white for 'Sunday'; Rouw, Case, Gosavi, & Ramachandran, 2014). Sound-symbolism correspondences such as the well-known Kiki-Bouba effect (Köhler, 1929; Ramachandran, 2001b), in which a star-shaped figure is highly associated with the name 'Kiki' whereas a rounded figure tends to be called 'Bouba', have also been theorised to resemble the properties of synaesthetic associations (Milán et al., 2013; see Brang & Ramachandran, 2020 for a general review).

At the same time, Newell (2013) found that synaesthetic colour experiences of grapheme-colour synaesthetes could be triggered by letters encoded through touch rather than the usual visual or auditory sensory modality channels. Related findings were observed in an investigation that assessed the McGurk illusion (McGurk and MacDonald, 1976) to investigate the emergence of synaesthetic colour perception in grapheme-colour synaesthetes (Bargary, Barnett, Mitchell, & Newell, 2009). The McGurk illusion is a well-established multisensory paradigm which occurs when different auditory and visual semantic stimuli that are shown simultaneously are perceived fused into a new percept (e.g. seeing someone vocalise ‘gate’ concurrent to the sound ‘bait’ being pronounced causes the illusory hearing perception of ‘date’ to the observer). Interestingly, the authors found that synaesthetes experienced qualitatively and quantitatively (in terms of colour spectral distance) different synaesthetic colours for the perception of the illusory words (i.e. ‘date’) and the auditory components alone (i.e. ‘bait’). Thus, synaesthetic pairings do not only resemble typical cross-modal associations, but these studies seem to suggest that synaesthetic perception is also influenced by cross-modal processes.

As a matter of fact, there is evidence suggesting strong overlaps between the neuronal mechanisms responsible for the integration of information coming from different sensory modalities (i.e. multisensory integration) and synaesthetic processing (e.g. Bankieris & Simner, 2015; Bien, ten Oever, Goebel, & Sack, 2012; Simner, Gärtner, & Taylor, 2011; Ward et al., 2006). In particular, specific areas of the parietal cortex have been suggested to play an influential role in both synaesthetic and non-synaesthetic multisensory integration processes. For example, increased connectivity between the intraparietal sulcus and the superior parietal lobule and early sensory areas has been observed to aid normal multisensory binding by facilitating reaction times to audio-visual stimuli in typically developed individuals (e.g. Brang, Taich, Hillyard, Grabowecky, & Ramachandran, 2013b). The importance of the parietal cortex in the integration on multisensory stimuli seems to be reinforced by other investigations that have confirmed the involvement of this brain area in the processing of other types of multisensory stimuli such as visuo-tactile (e.g. Pasalar, Ro, & Beauchamp, 2010) or audio-somatosensory (e.g. Foxe et al., 2000).

On the other hand, integration of synaesthetic associations has also been suggested to be mediated by the parietal lobe, which has been observed to be anatomically larger (specifically, the superior parietal lobule) for synaesthetes than controls (Hängii et al., 2008; Jancke et al., 2009; Weiss & Fink, 2008). Importantly, whole-brain analyses have also reported increased clusters of activation of the parietal cortex, near the intraparietal sulcus and in the angular gyrus, during the synaesthetic experience (Laeng et al., 2011; Nunn et al., 2002; Paulesu et al., 1995; Rouw & Scholte, 2010; Steven et al., 2006; Weiss et al., 2005). Supporting this, several studies have documented suppression or attenuation of synaesthetic inducer-concurrent experiences following transient disruption of parietal areas through transcranial magnetic stimulation (right posterior parietal lobe: Esterman, Verstynen, Ivry, & Robertson, 2006; Muggleton, Tsakanikos, Walsh, & Ward, 2007; parieto-occipital junction: Rothen et al., 2010).

However, the question remains: are the type of cross-modal associations that identify synaesthetic experiences indicative of a general pattern of increased multisensory integration in these individuals? A few studies have investigated this hypothesis through multisensory illusion paradigms, commonly used to assess the strength of cross-modal integration as they are considered highly automatic perceptual processes (e.g. Stevenson et al., 2014). In particular, two types of illusory paradigms have been mostly used to study multisensory integration in synaesthesia: the sound-induced flash illusion (SIFI; Shams, Kamitani, & Shimojo, 2000) and the McGurk illusion (McGurk & MacDonald, 1976). Both illusions arise when incompatible audio-visual information is presented approximately at the same time and from the same location. The brain treats simultaneous stimuli coming from nearby locations as belonging to the same object; i.e. resolves the conflict binding together stimuli that are in fact incongruent.

The SIFI (Shams et al., 2000), also called the double-flash illusion, occurs when a single flash of light is presented together with two beep sounds (1F2B), creating the false or illusory perception of seeing two flashes instead of one to the perceiver. This is the fission or canonical double-flash illusion. In addition, a weaker illusion, called the fusion illusion, has also been documented (Andersen, Tiippana, & Sams, 2004; Mishra, Martínez, & Hillyard, 2008; Shams, Ma, & Beierholm, 2005). In this case, two flashes are accompanied by one beep (2F1B) and the subject tends to see one flash instead of two. The integration or binding window of the SIFI has been established at a stimulus-onset asynchrony (SOA; i.e. amount of time between the start of each stimulus) of -/+ 150 ms (Brang, Williams, & Ramachandran, 2012; Foss-Feig et al., 2010; Neufeld, Sinke, Zedler, Emrich, & Szycik, 2012; Shams et al., 2000; Whittingham, McDonald, & Clifford, 2014), being optimal, especially for synaesthetes, at a -/+ 50-100 ms SOA (Neufeld et al., 2012). The percentage of perceived illusions (i.e. number of reported 2 flashes in the 1F2B fission illusion or 1 flashes in the 2F1B fusion illusion) is regarded as a measure of strength of audio-visual multisensory integration (Foss-Feig et al., 2010; Neufeld et al., 2012).

Brang and colleagues (2012) compared a group of 7 grapheme-colour synaesthetes and 25 controls in different experimental conditions. The authors observed that synaesthetes, compared to non-synaesthetes, significantly reported perceiving *more* double flashes in the fission illusion condition (1F2B; +50 ms SOA) than in the fission control condition (1F2B; +300 ms SOA) and thus concluded that synaesthetes showed increased multisensory processing. Although not reported as such, Brang et al. (2012) also assessed the fusion illusion (2F1B; +50 ms SOA) as a control condition, but they did not observe any group differences. Contrasting the results of Brang and colleagues, Neufeld et al. (2012) obtained the opposite findings in a study that assessed 18 grapheme-colour and/or auditory-visual synaesthetes and 22 controls on the same illusion paradigm. In particular, they evaluated the fission (1F2B; presented at different SOAs: from 25 to 500 ms before or after the first sound) and fusion (2F1B) illusions. Results showed that synaesthetes perceived significantly *fewer* fission illusions (1F2B) than controls at the specific binding windows of +50 ms and +100 ms SOA, and that no such group differences were found for a baseline 1F1B condition. No differences were observed either for the perception rates of the fusion illusion (2F1B). Lastly, Whittingham et al. (2014) presented 21 grapheme-colour and 1 sound-colour synaesthetes and 33 controls with one, two, or three flashes and one, two, or three beeps (all SOAs +60 ms), resulting in 9 possible combinations of audio-visual stimuli or experimental conditions that included the fission and fusion illusions. Synaesthetes reported slightly fewer perceived (fission and fusion) illusion than controls, but the results were statistically non-significant.

In line with these findings, Sinke et al. (2014) investigated multisensory integration in synaesthetes with another audio-visual illusion: the McGurk illusion (McGurk & MacDonald, 1976). As mentioned above, this illusion arises due to the fact that incongruent simultaneous visual and auditory linguistic stimuli are sometimes fused to a new percept. Nineteen grapheme-colour and/or auditory-visual synaesthetes and 24 non-synaesthete controls synaesthetes were compared, results showing that synaesthetes experienced significantly *fewer* fusion responses (i.e. diminished McGurk effect) and that their answers were driven mainly by the auditory information. Sinke and colleagues (2014) conducted a second experiment analysing speech comprehension in a noisy environment on a new sample of participants (14 grapheme-colour and/or auditory-visual synaesthetes and 14 non-synaesthetes). Participants were showed congruent visual and auditory articulatory movements under different environmental noise conditions. In accordance with the McGurk effect findings, synaesthetes, compared to non-synaesthetes, benefited *less* from the visual articulatory information– i.e. they showed less integration of the audio-visual stimuli than controls under specific noise conditions. Therefore, Sinke et al. (2014) concluded that synaesthetes showed reduced multisensory integration.

Although multisensory integration processes seem to be involved in the synaesthetic experience (Bargary et al., 2009; Newell, 2013) and there are strong neuronal overlaps between the cortical areas involved in both processes (see above), evidence pointing to a general pattern of increased multisensory integration for synaesthetes compared to non-synaesthetes is inconsistent to date. In fact, findings from multisensory illusion paradigms suggest that synaesthetes might experience reduced multisensory integration. Therefore, current data is insufficient to give a clear answer regarding the relationship between multisensory integration and synaesthesia.

2.1.2 The role of attention and its complex interplay with multisensory integration

Although the results of the studies on multisensory illusion might suggest diminished multisensory integration in synaesthetes, it is important to note that other interpretations are possible. Perceptual illusions have been typically used to measure the strength of multisensory integration (e.g. Stevenson et al., 2014), but participants are explicitly instructed to ignore one sensory modality whilst completely focusing on another one (e.g. SIFI: pay attention to the flashes and ignore the beeps). Thus, selective attentional mechanisms (i.e. responsible for the filtering of relevant from irrelevant stimuli in the environment) might be also at play. Despite the longstanding assumption that multisensory integration (and multisensory illusions) operated in an attention-free mode, multiple investigations have recently challenged this view (see Koelewijn, Bornkhorst, & Theeuwes, 2010; Macaluso et al., 2016; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010; and Tang, Wu, & Shen, 2016 for reviews). Broadly speaking, evidence seems to suggest that whereas early, automatic integration occurs with high-salient stimuli and in low-resource-competing contexts; low-salient stimuli and high-resource-competing contexts lead to late, non-automatic integration modulated by attentional influences (Talsma et al., 2010). Hence, multisensory integration would take place at multiple stages and would be controlled by dynamic modulation of attention depending on the available resources (i.e. parallel integration framework; Calvert & Thesen, 2004), and on other factors such as context (including observer goal and task) and priors (i.e. knowledge and expectations of the observer regarding the stimuli and their causes) (Macaluso et al., 2016).

Such inter-modulating relationships between multisensory integration and attention have been documented in relation to the SIFI. For instance, Mishra, Martínez, and Hillyard (2010) examined the effects of attention on event-related potentials (ERP) components previously associated with susceptibility to the SIFI (specifically, the early occipital-temporal PD110/PD120 components; Mishra et al., 2008). The authors presented simultaneous audio-visual stimuli in the left and right visual fields but asked participants to focus on either of them at a time and ignore the other one. They found that when the stimuli were ignored, the amplitude of these ERP components was significantly reduced compared to when it was attended. Thus, Mishra and colleagues (2010) suggested that endogenous spatial attention had a role in the occurrence of the illusory effect. In line with these observations, van der Stoep, van der Stigchel, and Nijboer (2015) investigated the influence of exogenous spatial orienting on the multisensory integration of audio-visual stimuli under detection and localisation tasks. Whereas in the localisation task the target stimuli could be presented either right or left of a central fixation point, in the detection task the stimuli always appeared in the centre. In both tasks, there was a decrease in multisensory integration when the audio-visual targets were exogenously attended relative to when they were not. But, importantly, the integration of audio-visual stimuli was specially diminished when space was relevant (i.e. location task) compared to when it was irrelevant (i.e. detection task). Therefore, the authors concluded that exogenous attention influences multisensory integration when spatial orienting is relevant.

Extending these findings, de Haas, Kanai, Jalkanen, and Rees (2012) examined structural brain differences associated with the SIFI. The researchers observed that higher susceptibility to the illusion was significantly and strongly correlated with low local grey matter volume in the early retinotopic visual cortex, further confirming the critical involvement of the occipital area in the perception of the SIFI (Mishra et al., 2008; 2010). Moreover, de Haas et al. (2012) discussed the possible interactions of the illusion with attentional mechanisms and proposed that susceptible individuals could be either allocating more attention to the spatial location of the multisensory stimuli or being less able to supress the auditory stimuli via top-down attention. On the other hand, Kamke, Vieth, Cottrell, and Mattingley (2012) investigated whether specific brain areas involved in selective attention modulated the multisensory processing of the SIFI using transcranial magnetic stimulation (TMS). The authors observed that disruption of the angular gyrus within the right parietal lobe reduced participants' susceptibility to the illusion, concluding that this region does not only contribute to the binding of audio-visual stimuli but also has a role in the perception of attended events.

Reduced illusory susceptibility due to attentional modulation has also been observed in the McGurk illusion. For example, Tiippana, Andersen, and Sams (2004) examined the McGurk illusion in two conditions: a baseline condition in which participants were asked to focus their attention on the talking face, and a distracting condition in which participants were told to ignore the face and attend a visual distractor (i.e. a leaf moving across the face). Results showed a significant weaker McGurk effect for the latter condition. Alsius and colleagues confirmed and extended these findings in a series of experiments (Alsius, Möttönen, Sams, Soto-Faraco, & Tiippana, 2014; Alsius, Navarra, & Soto-Faraco, 2007; Alsius, Navarra, Campbell, & Soto-Faraco, 2005). In a first experiment (Alsius et al., 2005), concurrent to the typical reporting of phonemes of the McGurk task, subjects had to detect specific line-drawn pictures or common sounds superimposed on the speech video. In addition, in Alsius et al. (2007) participants were asked to attend the monitor while placing their fingers on tactile tappers. The task consisted of repeating back verbally any words they heard on the speaker in addition to responding via foot pedal to specific tactile targets interspersed amongst a stream of tactile vents delivered through the tactile tappers on the fingers. In both studies, susceptibility to the illusion was severely diminished when participants were concurrently performing the unrelated visual, auditory, or tactile tasks compared to when participants performed the same tasks without the additional attentional demands (Alsius et al., 2007) or to participants who only performed the illusory task (Alsius et al., 2005).

These attention influences on early neural integration of vision and audition in speech were further confirmed through the examination of ERP recordings while participants performed a similar McGurk dual-task paradigm (Alsius et al., 2014). In this case, individuals were asked to identify auditory, visual, or audio-visual presented syllables whilst a rapid visual stream of line-drawing pictures was shown simultaneously; in the dual condition they had to monitor the pictures for certain repetitions, whereas they were told to ignore these instructions in the single condition. In accordance to previous studies, the authors found a weaker McGurk effect in the dual-task compared to the single condition. They also replicated previous findings that the latency of the early auditory N1/P2 ERP complex was reduced for audio-visual compared to auditory speech stimuli, ratifying these components as markers of (speech) multisensory integration (Baart, Stekelenburg, & Vroomen, 2014; Knowland, Mercure, Karmiloff-Smith, Dick, & Thomas, 2014; van Wassenhove, Grant, & Poeppel, 2005). However, Alsius et al. (2014) critically observed as well that this latency was diminished when attention was loaded (i.e. dual-task condition), suggesting that attention modulates early neural processing of audio-visual speech by weakening the integration between these two sensory modalities. Going one step further, in a series of studies, Morís Fernández and colleagues showed that the McGurk illusion activates conflict resolution brain areas (e.g. anterior cingulate cortex, language specific: inferior frontal gyrus) and thus argued that the illusion might be mediated by general-purpose conflict mechanisms (Morís Fernández, Macaluso, & Soto-Faraco, 2017; Morís Fernández, Torralba, & Soto-Faraco, 2018; Morís Fernández, Visser, Ventura-Campos, Ávila, & Soto-Faraco, 2015).

In sum, these studies indicate that although the SIFI and the McGurk illusion (and, perhaps, multisensory illusions in general) can be informative about the strength of multisensory integration processes, they cannot be considered fully automatic processes and, therefore, attention's modulation role must be considered. As pointed above, this has implications for the multisensory integration studies concerning synaesthesia. If attention mediates multisensory illusions, could the observed differences between synaesthetes and non-synaesthetes in these paradigms be better explained from an attention framework? More specifically, is it possible that enhanced filtering or selective abilities, rather than reduced multisensory integration, could describe the results found in illusion perception in synaesthetes?

2.1.3 Do synaesthetes experience atypical attention?

The study of the role of attention and awareness in synaesthetic perception has attracted a great deal of research (see Rich & Mattingley, 2013 for a review). Some investigations support the idea that synaesthetic associations are processed pre-attentively, in a bottom-up fashion (Hubbard, Arman, Ramachandran, & Boynton, 2005; Laeng, 2009; Laeng, Svartdal, & Oelmann, 2004; Palmeri, Blake, Marois, Fanery, & Whetshell, 2002; Ramachandran & Hubbard, 2001a; 2001b; Sagiv, Heer, & Robertson, 2006a; Ward, Jonas, Dienes, & Seth, 2009). However, this evidence primarily comes from visual search paradigms based on 'pop-out' effects (i.e. an item embedded into other items that it is detected due to its uniqueness; in this case, [synaesthetic] colour being the factor that stands out). For that reason, and together with the results of other studies in visual search tasks (Edquist, Rich, Brinkman, & Mattingley, 2006; Gheri, Chopping, & Morgan, 2008; Nijboer & van der Stigchel, 2009; Rothen & Meier, 2009), some authors suggest that these findings should be better understood as a synaesthetic advantage to reject distractors or to group items. This is further supported by other investigations measuring synaesthetic congruency effects (i.e. conflict between synaesthetic stimuli inducers that match or not a display colour; e.g. synaesthetic red letter 'A' coloured red or coloured blue). These studies have shown that synaesthetic percepts are only experienced when sufficient attentional resources are available to bring the inducers into awareness (e.g. Mattingley, Payne, & Rich, 2006; Mattingley, Rich, Yelland, & Bradshaw, 2001; Rich & Mattingley, 2005; 2010; Palmeri et al., 2002; Sagiv et al., 2006a). Therefore, this evidence demonstrates that attention has a modulatory role on attention in synaesthesia, via top-down mechanisms (Rich & Mattingley, 2013).

Despite the fact that synaesthetes frequently and unpredictably experience irrelevant and potentially distracting percepts, phenomenological reports seem to indicate that many synaesthetes do not typically view their synaesthetic associations as a source of cognitive distress or interference. For example, Rich et al. (2005) conducted a study on the implications of grapheme-colour synaesthesia and participant KP commented on this point: "It's kind of like looking at your own nose – if you try, you can see it clearly, but you don't walk around the whole time 'seeing' your nose" (see Day, 2005 for other phenomenological testimonies). This seems to suggest that, despite synaesthetes cannot help experiencing their inducer-concurrent associations, they are able to largely ignore or filter out concurrents whenever necessary. Therefore, selective attention might not be only crucial for synaesthetic binding but could also be a useful resource in the general synaesthetic experience. One open question, already pinpointed in the previous section, is whether this 'special' filtering ability might extend beyond synaesthetes specific inducer-concurrent associations. Or, in other words, whether synaesthetes experience general enhanced filtering or selective attention abilities (for synaesthetic and synaesthetic unrelated material).

Evidence in this respect is limited. The impact of task-irrelevant information on performance is typically measured in the lab with classic conflict tasks such as the Stroop task (Stroop, 1935), the Eriksen flanker task (Eriksen & Eriksen, 1974), or the Simon task (Simon & Wolf, 1963), providing a measure of participants' attentional abilities. Only a few studies have directly compared synaesthetes and non-synaesthetes on these paradigms (especially the Stroop task), obtaining contrasting results (Mattingley et al., 2001; 2006; Rouw, van Driel, Knip, & Ridderinkhof, 2013; Van der Veen, Aben, Smits, and Röder, 2014). In the classic or standard Stroop task (Stroop, 1935), subjects are presented with colour words which can be either congruently coloured (e.g. word "red" coloured red) or incongruently coloured (e.g. word "red" coloured blue) and asked to name the colours displayed (and not the words). Responses are typically slower and less accurate for the incongruent colour words than the congruent ones due to the interference caused by the irrelevant information contained in the incongruent words. The difference between these two types of trials is what is known as the [Stroop] congruency effect and it is considered a measure of participants' filtering or selective attention abilities. The larger the effect, the larger the interference and, thus, the weaker the individual's ability to filter out irrelevant information. Van der Veen et al. (2014) measured behavioural and blood-oxygen level dependent (BOLD) responses of 13 grapheme-colour synaesthetes and

15 matched controls in the classic Stroop task. The behavioural results showed that synaesthetes had smaller congruency effects than non-synaesthetes (in this particular study, congruency effects were expressed as reaction time differences between incongruent and neutral trials – i.e. words presented in four different colours). Accordingly, the imaging data revealed that synaesthetes showed stronger activation of the rostral cingulate zone (a brain area generally involved in the detection of interference; Carter & van Veen, 2007) for the neutral stimuli. Thus, the two subpopulations showed differences in stimuli conflict processing.

However, other investigations have failed to observe differences between synaesthetes and non-synaesthetes in the Stroop task or other classic conflict tasks. Mattingley and colleagues (2001; 2006) compared groups of grapheme-colour synaesthetes and non-synaesthetes (15 vs. 15 and 14 vs. 14, respectively) in the standard Stroop task to check for baseline differences in participants' susceptibility to interference processing. In both studies, the authors observed expected congruency effects with significantly slower reaction times for incongruent compared to congruent trials, but there were no overall differences between groups or interactions between congruency type and group. On the other hand, Rouw et al. (2013) conducted a series of experiments on executive functions in synaesthesia. First, they assessed two groups of 15 grapheme-colour synaesthetes and 15 controls in the classic Stroop task, the reversed Stroop task (i.e. respond to the meaning of the word and not to the colour displayed), and a Stroop switching task in which these two tasks were intermixed. Results showed clear and significant congruency effects for the classic and the reversed Stroop tasks, both when conducted independently and in the switching task. In addition, there was an interaction between task switching and congruency; i.e. congruency effects were more pronounced in the Stroop switching task than in the independent tasks. However, none of these interacted with group and no main effects of group were observed either for any of the tasks.

Amongst other tasks, in a second set of experiments Rouw and colleagues (2013) compared 20 grapheme-colour synaesthetes and 20 non-synaesthetes in another classic conflict task, the Eriksen flanker task (Eriksen & Eriksen, 1974). There are several variations of the flanker task, but in its basic version, participants are presented with a central target (e.g. an arrow pointing to the left) and are told to ignore irrelevant distractors or flankers presented at the periphery (e.g. additional arrows pointing to the left, congruent condition; or to the right, incongruent condition). Like in the Stroop or other conflict tasks, larger response interference is observed for incongruent compared to congruent trials. Similarly to the Stroop results of the same study, there were significant congruency effects, but these effects did not differ between groups. In addition, in Rouw et al.'s (2013) study, the flanker task was combined with a stop-signal task. This consisted of an auditory stop-signal (i.e. a tone) presented immediately after the stimuli on 25% of the trials in which participants were told to refrain from responding. Such instructions have been shown to further weaken inhibitory control (Ridderinkhof, Band, & Logan, 1999; Verbruggen, Liefooghe, Notebaert, & Vandierendonck, 2005). No differences were found either in this respect between synaesthetes and non-synaesthetes.

One possible reason for these contrasting results is that the classic conflict tasks used to assess selective attention in synaesthetes do not engage the same filtering mechanisms that allow synaesthetes to ignore their irrelevant concurrents. If the activation of synaesthetic concurrents engages multisensory processing pathways (see section 2.1.1), then the attentional filtering of these irrelevant percepts might particularly involve intermodal attention, which is responsible for the filtering of information coming from an irrelevant sensory modality. This type of attention also plays a critical role in multisensory illusions tasks (see section 2.1.2), as individuals are presented with simultaneous stimuli from different sensory modalities and are asked to focus in one modality and ignore the other – and synaesthetes have been shown to filter out illusory-inducing distractor information more efficiently than controls (see section 2.1.1). In addition, attention to location (spatial attention) and attention to stimuli from a specific sensory modality (intermodal attention) are mediated by different mechanisms with intermodal attention operating through selective modulation of modality-specific areas (e.g. Eimer et al., 1998; Macaluso, Frith, & Driver, 2002; Talsma & Kok, 2002). Thus, synaesthetes might have enhanced intermodal filtering abilities rather than general selective attention abilities.

2.1.4 The cross-modal congruency task: A paradigm to study attention and multisensory integration processes in synaesthesia

If synaesthetes are particularly efficient at filtering irrelevant distractors in a specific sensory modality while focusing on another modality, this advantage should be evident in tasks that are typically used to assess intermodal selective attention. The cross-modal congruency task (CCT; Spence et al., 1998) is a well-established and robust paradigm that has been used to investigate a variety of research topics, including, cross-modal exogenous spatial attention (e.g. Driver & Spence, 1998a; 1998b), temporal processing (e.g. Shore, Barnes, & Spence, 2006), distracter suppression (e.g. Marini, Chelazzi, & Maravita, 2013), multisensory interactions in peripersonal space (e.g. Maravita, Spence, & Driver, 2003; Spence, Pavani, Maravita, & Holmes, 2004b; 2008; van Elk, Forget, & Blanke, 2013), tool use (e.g. Maravita, Spence, Kennett, & Driver, 2002; Holmes, Calvert, & Spence, 2004; 2007; Sengül et al., 2013), the rubber hand illusion (e.g. Pavani, Spence, & Driver, 2000; Walton & Spence, 2004; Zopf, Savage, & Williams, 2010; 2013), or even the embodiment of a robotic prosthesis (Marini et al., 2014). The CCT has also been employed to measure multimodal interactions and response conflict in clinical conditions and other fields such as schizophrenia (e.g. Stekelenburg, Maes, van Gool, Sitskoorn, & Vroomen, 2013), autism spectrum disorder (e.g. Foss-Feig et al., 2010; Poole, Couth, Gowen, Warren, & Poliakoff, 2015), dyslexia (e.g. Facoetti et al., 2010), dyspraxia (e.g. Bair, Kiemel, Jeka, & Clark, 2012), brain-damaged patients (Spence, Kingstone, Shore, & Gazzaniga, 2001), or ageing (e.g. Poliakoff, Ashworth, Lowe, & Spence, 2006).

In its original version (e.g. Pavani et al., 2000; Spence, Pavani, & Driver, 2004a), the CCT is a visuo-tactile conflict task in which participants are asked to make speeded judgements regarding the elevation of a tactile target (i.e. vibration burst) presented to the index finger (top location) and thumb (bottom location) of either hand, whilst ignoring a concurrent visual flash presented close to the one of these top-bottom body locations. The visual and tactile stimuli are either shown at the same location (congruent trials; top flashes – index finger bursts or bottom flashes – thumb bursts) or at different locations (incongruent trials; top flashes – thumb bursts or bottom flashes – index finger bursts). Responses are faster and more accurate on congruent than incongruent trials giving rise to consistent congruency effects. Thus, CCT congruency effects can be considered a measure of the strength of intermodal selective attention abilities: the smaller the congruency effects, the stronger the capacity to filter out irrelevant stimuli in a second sensory modality. In addition, it is worth highlighting that the task-irrelevant sensory modality of the classic CCT is vision. This is of relevance because it matches the type of attentional distractors that a great number of synaesthetes (e.g. grapheme-colour synaesthetes) might naturally experience and, therefore, can be aimed at measuring similar cognitive components engaged during synaesthetic filtering.

A series of studies have investigated the proprieties of the CCT. Weaker congruency effects have been observed when the target-distractor sensory modalities are reversed (i.e. attend to visual targets and ignore tactile distractors; Walton & Spence, 1999; 2004), or when tactile distractors are paired with auditory targets (Merat, Spence, Lloyd, Withington, & McGlone, 1999). Attention has been shown to modulate the CCT. Several studies have observed that is harder to ignore irrelevant visual distractors when these are (exogenously) attended together with tactile targets at the same lateral location in space than when they are attended from different lateral spatial locations (e.g. same vs. different hands; e.g. Holmes, Sanabria, Calvert, & Spence, 2006; Marini, Romano, & Maravita, 2017; Spence et al., 2004a; 2004b). Making the target side or hand predictable (i.e. directing endogenous attention) has also been observed to facilitate the response latencies to the tactile targets, but this does not seem to affect the overall magnitude of congruency effects (e.g. Spence et al., 2004a; 2004b).

Besides attention, two other underlying mechanisms have been suggested to contribute to the congruency effects of the CCT: multisensory integration and response conflict

(Driver & Spence, 1998a; 1998b; Spence et al., 2004a; 2004b; Shore et al., 2006; Forster & Pavone, 2008; Holmes, 2012). Marini et al. (2017) directly addressed this issue assessing participants with the original task and a modified task with ad-hoc changes aimed at evaluating the role of attention and response conflict. In addition, participant's posture (i.e. hands) was also manipulated (focus on hand-mediated attentional binding). The results consistently showed that the more the spatial disparity between the visual and the tactile stimuli (i.e. targets and distractors), the larger the congruency effects. This would indicate a predominant role of response conflict for the congruency effects of the CCT, in line with the observations of Spence et al. (2004a; 2004b). However, it should be noted that weaker contributions from multisensory integration, in the absence of response conflict, and from hand-mediated attention binding, with modified posture and in the presence of response conflict, were also found; the authors concluding that a multifactorial interpretation might be the most accurate approach.

The CCT paradigm offers thus a unique framework to address the question of interest of this Chapter. That is, to determine whether synaesthetes are particularly better at ignoring irrelevant distractors in a different sensory modality (i.e. if they have enhanced intermodal attention abilities), providing an alternative explanation to previous findings of synaesthetic reduced susceptibility to multisensory illusory perception as well (see section 2.1.1). The following sections will investigate this and several follow-up questions using the classic version of the CCT and ad-hoc modified tasks. Study 1 and Study 2 will compare unimodal and cross-modal differences when vision is, respectively, the distractor and the target sensory modality. Study 3 will investigate different pairings of sensory modalities, evaluating visuo-tactile vs. audio-visual differences and vision acting both as the distractor and the target. Vision has a prominent role in our investigation because all the synaesthetes assessed throughout these studies experienced types of synaesthesias that involved visual concurrents (mostly, grapheme-colour and/or sequence-space synaesthesias). These synaesthetes were purposely screened in an attempt to keep the synaesthetic- and task-irrelevant distractors as compatible as possible, a condition necessary to evaluate the generalisation of synaesthetic filtering processes to synaesthetic unrelated stimuli.

2.2 Unimodal vs. cross-modal differences when vision is distractor (Study 1)

2.2.1 Introduction

Veen et al. (2014) found that synaesthetes experienced less response interference than non-synaesthete in the Stroop task (Stroop, 1935), a classic conflict task. However, other investigations have failed to observe such differences in the same or similar tasks (Mattingley et al., 2001; 2006; Rouw et al., 2013). Given the cross-modal nature of synaesthetic associations, it might be possible that, rather than differences in general selective attention abilities, synaesthetes might particularly present enhanced intermodal filtering abilities, which specifically filter information coming from an irrelevant sensory modality. As a matter of fact, synaesthetes have been found to show reduced susceptibility to multisensory illusion perception (in which attention is required to focus in one sensory modality while ignoring the other) in some studies (Neufeld et al., 2012; Sinke et al. 2014) but not on others (Brang et al., 2012 and Whittingham et al., 2014). One possible explanation for the inconsistency in these results is that the illusory stimuli present special properties and do not tap into the same cross-modal mechanisms involved in synaesthetic processing. For that reason, in this study we propose a new paradigm to assess whether synaesthetes present enhanced intermodal attention abilities.

To do this, we compared the performance of a group of synaesthetes to that of a matched group of non-synaesthetes in the cross-modal congruency task (CCT; Pavani et al., 2000; Spence et al., 2004b), a well-known paradigm that measures intermodal selective attention. In this task, participants are asked to respond to tactile targets while ignoring simultaneous visual targets (see section 2.1.4 for complete details). In some trials, the target and the distractor stimuli are presented in the same location (e.g. top/top; congruent condition) and in other in opposite locations (e.g. top/bottom; incongruent condition). Responses are typically slower and more prone to errors on incongruent than congruent trials. These differences in reaction times and accuracy rates are known as congruency effects (CE) and they are considered a measure of strength of intermodal selective attention abilities: the smaller the CE, the stronger the ability to ignore irrelevant stimuli in a second sensory modality.

To keep the congruency task as close as possible to the type of attentional filtering synaesthetes might experience in their inducer-concurrent associations, we matched the sensory modality of the task-irrelevant distractors in the CCT (i.e. vision) with the sensory modality of synaesthetes' concurrents. Since the most common forms of synaesthesia involve visual concurrents, only synaesthetes with at least one synaesthesia type involving vision as the concurrent modality (e.g. synaesthetes with grapheme-colour synaesthesia or sequence-space synaesthesia; hereafter referred as -visual synaesthetes) were included in this study. The CCT was therefore aimed at measuring similar cognitive mechanisms involved during synaesthetic filtering. We predicted that if synaesthetes' constant need to disregard their automatic and irrelevant synaesthetic associations is generalised to other non-synaesthetic multimodal stimuli, they should show smaller CE in the CCT compared to non-synaesthetes, reflecting enhanced intermodal selective attentional abilities.

In addition, to confirm the lack of differences between synaesthetes and non-synaesthetes with respect to general filtering skills, we asked participants to perform the Eriksen flanker task (FT; Eriksen and Eriksen, 1974). We chose this task over the Stroop, another well-known classic conflict task used in synaesthetes in previous studies, in order to avoid synaesthetic interference confounds due to the colour stimuli used in this task. The FT is a well-established paradigm that has been widely used to assess distractor inhibition and response competition (see Eriksen, 1995 for an historical review) and a variation of this task is part of the attention network task (ANT), which is routinely used to measure the executive control network of attention in developmental and clinical settings (e.g. MacLeod et al., 2010). Participants are typically asked to make speeded choice responses to a central target whilst ignoring the irrelevant distractors (flankers) presented at the periphery. Targets and distractors are mapped to congruent and incongruent conditions and, similarly to the CCT, differences in mean reaction times and error rates between incongruent and congruent trials (i.e. CE) reveal the difficulty to ignore the irrelevant distractors and thus the strength of participants' filtering abilities. If the cognitive mechanisms activated during the management of irrelevant synaesthetic sensations and irrelevant information in the FT are at least partially overlapping, synaesthetes should also show an advantage at distractor filtering in this task.

2.2.2 Methods

2.2.2.1 Participants.

A total of 52 subjects participated in this study. All participants reported no known neurological illnesses and normal or corrected-to-normal vision. The study was approved by The University of Edinburgh's Psychology Research Ethics Committee and followed the ethical guidelines laid down in the Helsinki Declaration. Participants were recruited via the University's employment website and convenience sampling, and they received a small monetary compensation (£7-12). Informed consent was obtained from all participants.

2.2.2.1.1 Synaesthesia screening and classification of participants.

We devised a thorough methodology to determine the synaesthetic status of our participants. Participants were divided into synaesthetes and non-synaesthetes following the completion of an ad-hoc synaesthesia screening interview, the Edinburgh Synaesthesia Screening Assessment (ESSA), that we developed partially adapting from Banissy et al. (2009) and Kusnir and Thut (2012). The interview thoroughly explored a large number of types of synaesthesia (taking Day's, 2019 register as reference) and inquired about the frequency, constancy, location, and stability people self-reportedly experienced each type of synaesthetic association (see Chapter IV for further details and Appendix A for a copy of the instrument).

Participants who reported some type of grapheme-colour or sound-colour synaesthesia further completed the Synesthesia Battery (SB) (Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007). The SB is a standardised battery which measures the internal consistency of synaesthetic experiences related with -colour [i.e. as the concurrent] for several triggers including letters, numbers, weekdays, months, piano scale notes, chords, and instruments. In this test, subjects are presented with a single letter at a time on a computer screen and they must choose the colour that best matches their synaesthetic experience from a colour palette of over 16 million colour bitmaps. Each letter is presented three times in a randomised order. Then, the spectral distance between the different colour choices for each letter is calculated and averaged. The less the distance between colours, the stronger the synaesthetic association is regarded, scores below 1.0 traditionally being considered to indicate the presence of synaesthesia and scores of 1.0 or above absence of it. However, some authors argue that a cut-off score of 1.43 is more discriminant between synaesthetes and non-synaesthetes and is used in some studies as an alternative, looser threshold (e.g. Carmichael et al., 2015; Rothen, Seth, Witzel, & Ward, 2013; Ward et al., 2017a). These two thresholds allow thus two distinguish two types of synaesthetes: ‘strong’ (those who pass the consistency test at the strict criterium; <1.0) and ‘weak’ (those who pass it at the loose criterium; <1.43). If a person took more than one test, the minimum average score obtained in any of them was taken as the reference score for classification purposes.

Following the methodological procedures of previous studies (e.g. Havlik et al., 2015; Price, 2009; Rizza & Price, 2012), participants who responded ‘Yes’ to the interview question tapping into spatial-sequences (“Do you see any of the following items as being arranged in specific patterns in space? I.e. the alphabet, the days of the week, the months, the numbers, the musical notes, and/or other”), were further prompted to describe (i.e. “How often do you see it?”, “Does the arrangement always have the same pattern?”, “Where do you see it?”, “When did you start seeing it?”, etc.) and draw their sequences. In addition, they were also asked about the locus of their synaesthetic perceptions to classify them into ‘associators’ (i.e.

perception of the synaesthetic sensation into their ‘mind’s eye’) or ‘projectors’ (i.e. synaesthetes who experience these sensations outside their body; Dixon et al., 2004). To do that, we used relevant questions adapted from previous studies (e.g. Rouw & Scholte, 2007; Skelton, Ludwing, & Mohr, 2009): e.g. “Do you see the colours superimposed on the letters? Or are the letters not coloured, but you are aware that they have specific associated colours?”.

People that responded ‘Never’ to all the possible types of synaesthesias were classified as non-synaesthetes. We considered synaesthetes those participants who expressed having any synaesthetic experiences on a highly regular basis (i.e. answer option ‘Always’ of the frequency aspect of the interview), as this is indicative of highly automatic experiences, one of the main criteria established in the literature (e.g. see Ward, 2013 for a review). In addition, if they completed any consistency tests for -colour experiences, they had to pass it at the strict threshold. If participants only experienced [-colour testable] synaesthesias and failed the tests, they were classified as non-synaesthetes; if they passed it just at the loose threshold, they were classified as ‘weak’ synaesthetes. If participants experienced other types of synaesthesias besides the testable ones only in a ‘Sometimes’ basis and failed the test, they were also classified as ‘weak’ synaesthetes; if they experienced them on a highly regular basis (i.e. ‘Always’), failing on the tests was not a condition to strip them of their (strong) synaesthetic status. Lastly, there were some people who only reported having synaesthetic experiences ‘Sometimes’ and which were tentatively grouped into the ‘weak’ synaesthetes group as well. If they completed any consistency tests and passed them at the strict criterium, they were classified as (strong) synaesthetes; if they failed them (at both criteria) and did not experience any other synaesthesias, they were classified as non-synaesthetes.

2.2.2.1.2 Final sample.

Of the 52 people who completed the screening procedure and behavioural tasks, 34 were included in the final sample: 16 synaesthetes and 18 age-matched non-synaesthetes (demographics are reported in Table 1)[1]. All synaesthetes experienced at least one synaesthesia type involving vision as the concurrent modality; two people were excluded from the study for failing this criterion. We also removed sixteen people who were classified as 'weak' synaesthetes (all of them experiencing at least one -visual synaesthesia type), as we considered that their synaesthetic (or non-synaesthetic) status was unclear and that differences between synaesthetes and non-synaesthetes should emerge in the first instance with strong synaesthetes.

Table 1.
Descriptive, chi-square (χ^2), and t*-statistics of Study 1 groups' demographics.*

Demographics	Synaesthetes	Non-synaesthetes	Statistics
N (male)	14 (2)	14 (4)	$\chi^2(1) = .55$, $p = .46$
Age (SD)	25.7 (2.77)	24.1 (2.34)	$t(32) = 1.86$, $p = .072$
Handedness (left)	15 (1)	18	$\chi^2(1) = 1.16$, $p = .28$
Nº of (native) languages* (SD)	1.13 (.34)	1.22 (.55)	$t(32) = .61$, $p = .545$
Level of education** (SD)	3 (.73)	2.83 (.62)	$t(32) = .72$, $p = .48$

N = Sample size, SD = Standard Deviation.
* Nº of (native) languages: 1 = Monolingual, 2 = Bilingual, 3 = Polylingual.
** Level of education: 1 = High School, 2 = Undergraduate, 3 = Master, 4 = PhD, 5 = Postdoc.

Almost all synaesthetes reported multiple types of synaesthesias, with an overall average of 10 types (range: 2-20). Sixty-three percent of them experienced synaesthesias related with -colour and 69% of them sequence-space synaesthesias. Sixty percent of them also experienced ticker-tape synaesthesia (i.e. seeing spoken words or thoughts as 'subtitles'; e.g. Chun and Hupé, 2013), 44% mirror synaesthesias (i.e. experiencing tactile sensations in response to other people being touched or getting hurt; e.g. Ward & Banissy, 2015), or 44% ordinal-linguistic personification synaesthesias (i.e. attribution of personalities and/or genders to linguistic sequences such as numbers or letters; Simner & Holenstein, 2007).

[1] Given that there were no direct prior studies available, power analyses were performed to assess sample size taking the publications of Neufeld et al. (2012) and Sinke et al. (2014) as references. The combined results determined that a sample size of total *N* = ±34 (±15 synaesthetes and ±19 controls) was associated with an error probability α = .05 and power = .50. Thus, our proposed sample of 16 synaesthetes and 18 controls should be adequate for the aims of this study. The analyses were conducted with GPower 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007).

Synaesthetes who completed -colour consistency tests obtained, on average, a score of .61 points (SD = .33). The sequence-space descriptions and drawings were analysed in detail and their phenomenology was established consistent with the proprieties of synaesthesia in general and of this type of synaesthesia in particular (e.g. Cytowic, 2002; Price & Mentzoni, 2008; Sagiv et al., 2006b). We looked into things such as the richness or detail in describing the experience, the properties of the experience (e.g. paying special attention to the presence of characteristic traits like oval, staircase-alike, or other non-lineal shapes; 2 or 3-dimensionality; the capacity to zoom in and out or move around the visualised sequence; relative position from the body; etc), or the way of experiencing them (i.e. if they are automatic or not – e.g. affirmative responses to questions like "If I ask you what are you doing next Monday, does the image of the shape pop immediately and involuntarily?"). All synaesthetes were classified as 'associators'.

2.2.2.2 Experimental procedure.

The study took place in a dimly lit, sound attenuated room. Participants sat in a comfortable chair and rested their heads in a chinrest to maintain a constant distance from the stimuli displays. Stimuli presentation for both tasks was controlled and responses were recorded via E-Prime 2.0® software and hardware (Serial Response Box 200A®, Psychology Software Tools). Each participant performed the two tasks: the cross-modal congruency task (CCT) and the flanker task (FT). The order of the tasks, as well as the stimulus-to-response mapping for the CCT task (see below), were counterbalanced between participants. Before the beginning of each task, participants completed a practice block (12 trials) which was repeated if necessary. The study lasted approximately 60 minutes.

The CCT task was based on Pavani et al.'s (2000) and Spence et al.'s (2004b) studies. A black rectangular cuboid (70 x 35 x 35 mm) was positioned on the table in front of the participants (23 cm from the table edge where the chinrest was attached) and aligned with their body midline. Participants held the cuboid with the index finger and thumb of their dominant hand (placed on the top and bottom ends of the cuboid, respectively). Two tappers used to deliver the tactile targets were attached to the participants' hand, one to the index finger and one to the thumb (Miniature Solenoid Tappers-3 and Miniature Solenoid Controller-3.4® hardware, Mechanical & Electronic Solve). To mask the sound of the tappers, white noise (44.1 kHz frequency) was presented via headphones throughout the task at 60 dB(A). Two LED lights (diameter = 2 mm), used to present the visual distractors, were attached to the cuboid, one at the top and one at the bottom, next to the participants' fingers (controlled via Heijo Basic Visual Controller 291VISB® hardware, Heijo Research Electronics). A white pin at the centre of the cuboid served as fixation point.

Tactile and visual stimuli were presented in this task. On each trial, a tactile target was presented either to the top or bottom finger and consisted of three 50 ms onset periods during which a rod made contact with the skin, interleaved by two 50 ms offset periods. The visual distractor (illumination of the top or bottom LED) consisted of three successive 50 ms green flashes separated by two 50 ms offset periods (250 ms total duration). Each trial started with the presentation of the stimuli (250 ms), followed by a 1,550 ms empty interval in which responses were collected (total response window of 1,800 ms following stimulus onset), and by a variable inter-trial interval (ITI) randomly selected between 100-500 ms. Three different types of trials were presented: congruent, incongruent, and neutral. The tactile target and the visual distractor were simultaneously presented from the same location (top or bottom) on congruent trials, and from opposite locations (tactile stimulus top and visual stimulus bottom location, or vice-versa) on incongruent trials. On neutral trials, only the tactile target was presented (top or bottom location) (Fig. 1).

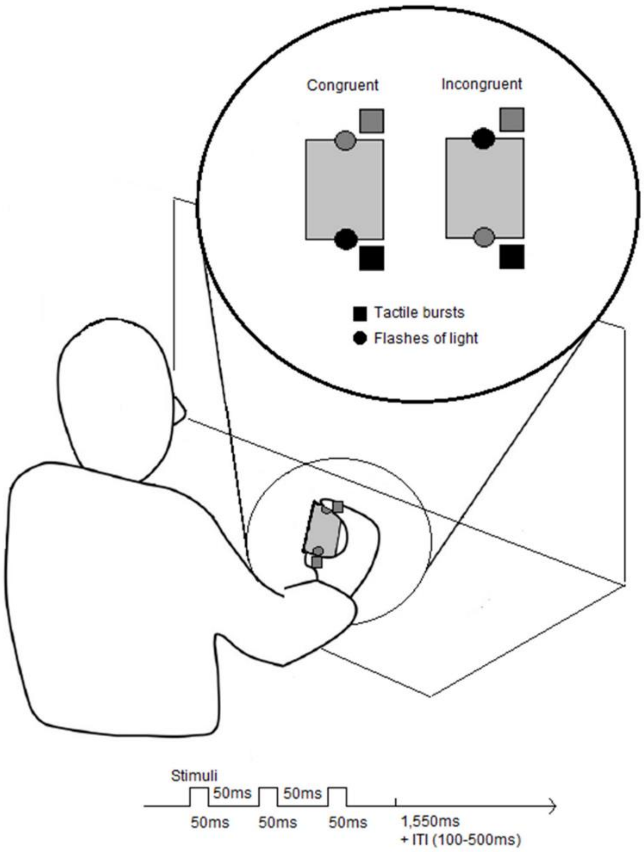


Figure 1. Display, type of trials, and timeline of Study 1 Cross-modal Congruency Task (adapted from Pavani et al., 2000 and Spence et al., 2004b); ITI = Inter-trial interval.

Regarding the FT, participants were instructed to perform an elevation discrimination task reporting via pedal press the location (top/bottom) of the tactile targets while ignoring the visual distractors when present. Half of the participants had to press the left pedal with their toes to indicate top location and the right pedal with their heel to indicate bottom location, and the other half followed the opposite mapping. Participants were also instructed to continuously keep their gaze on the fixation point and to answer as rapidly and accurately as possible. Participants completed three experimental blocks of 96 trials. Within each block, congruent, incongruent, and neutral trials were equally likely (32 trials per type) and randomly intermixed.

The FT experimental task was based on Eriksen and Eriksen (1974) design. Visual stimuli were presented on a computer monitor situated at a distance of 100 cm from the participant and consisted of black arrows (pointing left and right) and diamonds of 11.5 x 11.5

mm (0.66° of visual angle) on a light grey background. The centrally presented left or right arrow (target) was flanked by two additional stimuli on each side (distractors). Distractors were diamonds on neutral trials, whereas they were left or right arrows on congruent and incongruent trials pointing to the same or opposite direction, respectively, as indicated by the target. Each trial started with the presentation of a central fixation cross (6 x 6 mm black cross) for a duration randomly selected between 300-500 ms, followed by the display of the stimulus array for 100 ms. There was a total response window of 1,000 ms following stimulus onset and a variable ITI randomly selected between 500-700 ms (Fig. 2).

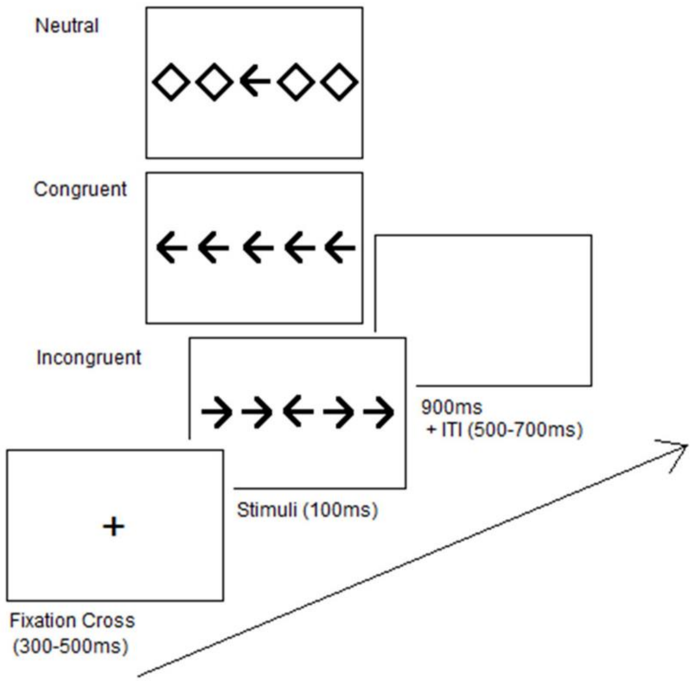


Figure 2. Type of trials and timeline of Study 1 Flanker Task (adapted from Eriksen & Eriksen, 1974); ITI = Inter-trial interval.

Participants were instructed to report via button press (keys 1 and 2 of the Serial Response Box, operated by the left and right index fingers) the direction (left vs. right) of the target (central arrow) while ignoring the distractors (flanking stimuli). Participants completed 3 blocks of 96 trials and within each experimental block, congruent, incongruent, and neutral trials were equally likely (32 trials per type) and presented in a randomised order.

2.2.2.3 Data analyses.

Separate analyses were conducted on error rates (ER) and reaction times (RT). Responses with RT exceeding ±3 standard deviations from the mean (calculated separately for each participant, task, and type of trial[2]; e.g. Igarashi, Kitagawa, Spence, & Ichihara, 2007) were considered as outliers and excluded from both ER and RT analyses. In the error rates analyses, ER reflected the percentage of correct responses after removal of omissions (i.e. no-response trials) and outliers. For the RT analyses, mean responses were calculated excluding choice-errors as well. Any participants with mean ER above 50% for any task condition were excluded from both types of analyses.

First, we conducted mixed analyses of variance (ANOVAs) with ‘Trial type’ (neutral, congruent, incongruent) as the within-subjects factor and ‘Group’ (non-synaesthetes, synaesthetes) as the between-subjects factor, separately for each task. Further pairwise comparisons and independent *t*-tests were carried out as appropriate following significant effects. Whenever necessary, *p*-values were adjusted for multiple comparisons with Bonferroni correction, and the Greenhouse-Geisser estimates of sphericity were used to report the results of the mixed ANOVAs when Mauchly’s test indicated that the assumption of sphericity had been violated. The analyses were conducted in Jamovi 0.9 (Jamovi Project, 2018) and SPSS 24 (IBM Corporation, 2016).

To assess the strength of the results observed, we then performed Bayesian interference analyses, a probabilistic approach which expresses the degree of belief in an event in terms of amount of evidence in favour of the null or the alternative hypotheses (e.g. Lee & Wagenmakers, 2005). Bayesian analyses provide a series of advantages in contrast to traditional *p*-value null hypothesis significance testing. First, they provide evidence for the null hypothesis as well as the alternative hypothesis. Second, Bayesian hypothesis testing is not

[2] Due to technical problems with the stimuli presentation software, only trials in which a left-pointing target was presented could be included in the analyses (differences between left and right trials were previously checked and rejected).

affected by the sampling plan (i.e. each added participant contributes towards the desired level of certainty), preventing null hypothesis testing' failure to detect effects due to small sample sizes or avoiding to observe effects in small samples that may be difficult to replicate. Lastly, Bayesian results fall into a continuous scale that quantifies the degree of certainty or probability of the observed evidence, as opposed to the dichotomous logic of null hypothesis testing (i.e. support for the alternative hypothesis or not). See Krypotos, Blanken, Arnaudova, Matzke, & Beckers, 2017 for an in-depth discussion about the topic.

In order to simplify the full model, we conducted separate ER and RT Bayesian mixed model ANOVAs for each task using default prior scales and the variable 'Congruency' (congruent, incongruent[3]) as the within-subjects factor and 'Group' (non-synaesthetes, synaesthetes) as the between-subjects factor. For each analysis, four models were compared to the null model: the 'Congruency'-only model, the 'Group'-only model, the non-interaction model ('Congruency' + 'Group'), and the interaction model ('Congruency x 'Group). We report the Bayes Factors (BF_{10}) associated with each model. In addition, we examined the inclusion effects of the factor 'Group' and the interaction term (Inclusion BF_{10}). Following Jeffreys (1961), a BF_{10} of 1-3 provides no evidence or anecdotal evidence for the alternative hypothesis, 3-10 moderate or substantial evidence, 10-30 strong evidence, 30-100 very strong evidence, and >100 decisive or extreme evidence. Analogously, a BF_{10} of 1-.33 offers anecdotal evidence for the null hypothesis, .33-.10 substantial evidence, .10-.03 strong evidence, .03-.01 very strong evidence, and <.01 decisive evidence. The analyses were conducted in Jasp 0.9 (JASP Team, 2018).

Lastly, we conducted linear mixed model analyses. In comparison to mixed ANOVAs, linear mixed models are more powerful and flexible (Armstrong, 2017). They are useful for complex models with multiple number of factors (as long as the sample sizes are large enough). Linear mixed models do not require complete or balanced data, have more lenient

[3] Neutral trials were omitted both from the Bayesian interference and the multivariate mixed model analyses given that no effects for this type of trials were observed in the main analyses (section 2.2.3.1) and in order to focus on the congruency aspect of interest.

assumptions, and show increased power to detect effects (Ma, Mazumdar, & Memtsoudis, 2012). Last but not least, these models they allow to measure the influence of participants' individual variability through random effects or effects that vary from participant to participant as opposed to fixed effects or effects that are constant across individuals (e.g. Gardiner, Luo, & Roman, 2008).

We run linear mixed model analyses of accuracy rates (AR) and RT as a function of the interaction between 'Congruency' (congruent, incongruent; see Footnote 3) and 'Group' (non-synaesthetes, synaesthetes), separately for each task. The models also included the random effect of participant with random slopes for type of trial (congruent, incongruent). Due to the non-normality of the AR data (see Footnote 4), we specifically performed generalised linear mixed models applying the Laplace Approximation method to estimate degrees of freedom and generate *p*-values for these analyses. We run linear mixed models fitting with Maximum Likelihood Estimation and Satterthwaite's method for the RT analyses. It should be noted that these analyses are computed on single data points rather than averaged data, therefore we did not apply outlier filters or removed underperforming participants; only omissions were excluded due to their uninformative and bias risk nature.

All analyses were performed in R 3.5.1 (R Core Team, 2018) with the following packages: *lme4* (Bates, Maechler, Bolker, & Walker, 2015), *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017), and *nlme* (Pinheiro, Bates, Debroy, Sarkar, & R Core Team, 2018). In addition, the packages *Matrix* (Bates & Maechler, 2018), *MuMIn* (Barton, 2018), *car* (Fox & Weisberg, 2011), *carData* (Fox, Weisberg, & Price, 2019), *jtools* (Long, 2019), and *ggplot2* (Wickham, 2016).

2.2.3 Results

2.2.3.1 Mixed Analyses of Variance (mixed ANOVA).

Outliers and omissions were excluded from the Cross-modal Congruency Task (CCT) analyses (3.10% of the total trials) and, overall, low ER were observed in both groups (M =

3.24, SD = 4.16 for synaesthetes and M = 3.43, SD = 3.38 for non-synaesthetes). The main effect of 'Trial type' (F(1.04, 33.25) = 21.4, $p < .001$, $\eta_p^2 = .40$) indicated the presence of higher ER on incongruent (M = 8.34, SD = 9.83) than congruent (M = .63, SD = 1.19; t(33) = 4.925, $p < .001$, d = 1.11) or neutral trials (M = 1.01, SD = 1.37; t(33) = 4.56, $p < .001$, d = 1.05), and no differences were found between congruent and neutral trials (t(33) = 1.48, p = .15) (Bonferroni-adjusted α = .017). The interaction between 'Trial type' and 'Group' was not significant (F(1.04, 33.25) = .17, p = .69) (see Fig. 3A).[4]

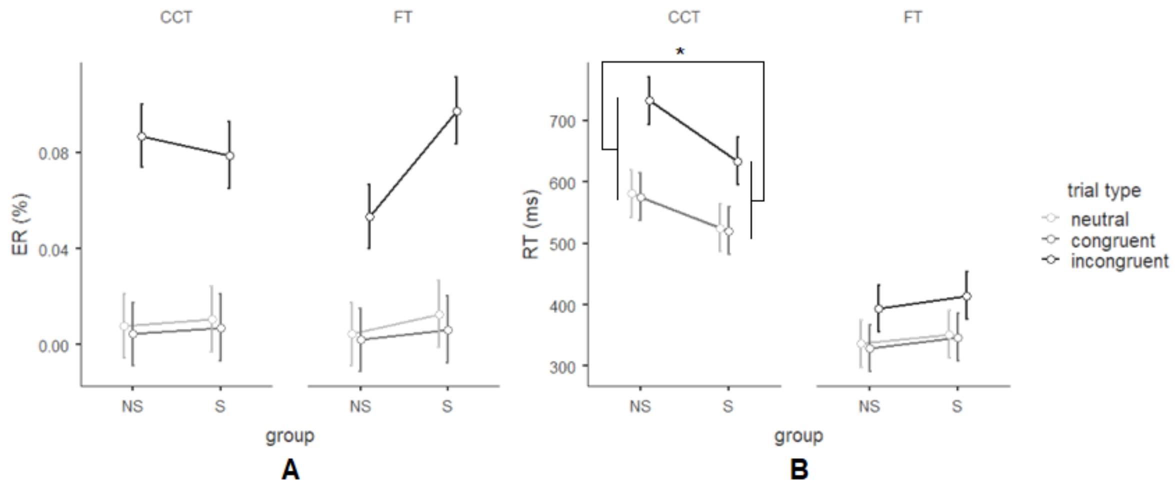


Figure 3. Mean error rates in percentages (ER; Figure **A**) and mean reaction times in milliseconds (RT; Figure **B**) and their corresponding standard error means (error bars) for neutral, congruent, and incongruent trials, for each task of Study 1 (Cross-modal Congruency Task – CCT; Flanker Task – FT) and group (non-synaesthetes – NS; synaesthetes – S). In the CCT, both the ER and RT analyses revealed a main effect of the factor 'Trial type' ($p < .001$), evidencing typical congruency effects (CE; i.e. incongruent trials presented slower RT and higher ER than congruent – and neutral – trials). There was also an interaction between 'Trial type' and 'Group' (p = .019) in the RT analysis. Specifically, synaesthetes presented faster RT for incongruent trials compared to non-synaesthetes. The determination of a significant smaller CE for synaesthetes (p = .020) confirmed this difference. Regarding the FT, the ER and RT analyses only showed a main effect of the factor 'Trial type' (both $p < .001$), reflecting the presence of CE.

The RT analysis revealed a main effect for the factor 'Group' (F(1, 32) = 4.26, p = .047), indicating that, overall, synaesthetes were faster than non-synaesthetes (M = 555, SD = 102 and M = 633, SD = 118, respectively; d = .71). As expected, significant differences emerged also between trial types (main effect of 'Trial type', F(1.16, 37) = 118, $p < .001$, η_p^2 =

[4] Since the ER variables violated the assumption of normality (as assessed by Shapiro-Wilk), we run an additional non-parametric Friedman χ^2 test for the repeated-measures analysis and Wilcoxon signed-rank Z tests for the pairwise comparisons. The analyses confirmed both the main effect of 'Trial type' (χ^2(2) = 46.3, $p < .001$) and the results observed for the paired-samples comparisons (incongruent vs. congruent trials: Z(33) = 1, $p < .001$; incongruent vs. neutral trials: Z(33) = 2, $p < .001$; neutral vs. congruent trials: Z(33) = 132, p = .33 – Bonferroni-adjusted α = .017).

.79). Follow-up pairwise comparisons (Bonferroni-adjusted α = .017) showed that RT were significantly slower for incongruent (M = 686, SD = 148) than for congruent (M = 550, SD = 108; $t(33) = 11.1$, $p < .001$; $d = 1.05$) and neutral trials (M = 553, SD = 104; $t(33) = 10$, $p < .001$; $d = 1.04$). No significant differences were found between congruent and neutral trials ($t(33) = .77$, $p = .45$).

Importantly, a significant interaction was observed between 'Trial type' and 'Group' ($F(1.16, 37) = 5.60$, $p = .019$, $\eta_p^2 = .15$). Follow-up independent *t*-tests conducted separately for each type of trial revealed significant differences between synaesthetes and non-synaesthetes on incongruent trials (M = 624, SD = 115 and M = 741, SD = 155, respectively; $t(32) = 2.48$, $p = .019$; $d = .86$), but not on congruent (M = 518, SD = 103 and M = 579, SD = 107, respectively; $t(32) = 1.69$, $p = .10$) or neutral trials (M = 523, SD = 100 and M = 580, SD = 103, respectively; $t(32) = 1.65$, $p = .11$) (Fig. 3B)[5]. To further investigate this finding and to quantify our effect of interest, the congruency effect (CE) (i.e. incongruent minus congruent trials average RT) was calculated for each subject. This measure was then submitted to a separate one-way ANOVA with 'Group' as between-subjects factor. The CE analysis showed a main effect of 'Group' ($F(1, 32) = 6$, $p = .020$, $\eta_p^2 = .16$), revealing a reduced CE in synaesthetes compared to non-synaesthetes (M = 106, SD = 60.2 and M = 162, SD = 71.4, respectively; $d = .85$) (Fig. 3B).

In the Flanker Task (FT), omissions and outliers represented a 1.41% of the total trials and overall ER were low for both groups: M = 3.94, SD = 4.63 for synaesthetes and M = 2.03, SD = 2.06 for non-synaesthetes. The main effect of 'Trial type' ($F(1.04, 33.2) = 20.8$, $p < .001$,

[5] Following a reviewer's suggestion on the published version on which this study is based (Mas-Casadesús & Gherri, 2017), an additional mixed ANOVA for the median RT of the CCT was carried out to confirm the robustness of the effects observed for the mean RT analyses across different measures of central tendency. The results revealed an analogous interaction between 'Trial type' and 'Group' ($F(1.09, 34.9) = 6.41$, $p = .014$, $\eta_p^2 = .17$), equally driven by the faster RT for synaesthetes compared to non-synaesthetes on incongruent trials (M = 611, SD = 113 and M = 734, SD = 151, respectively; $t(32) = 2.67$, $p = .012$; $d = .92$). The main effects of 'Trial type' and 'Group' were also ratified ($F(1.09, 34.9) = 107$, $p < .001$, $\eta_p^2 = .77$ and $F(1, 32) = 4.425$, $p = .043$, $\eta_p^2 = .12$, respectively). Specifically, incongruent trials (M = 676, SD = 146) were slower than both congruent (M = 535, SD = 107; $t(33) = 10.2$, $p < .001$; $d = 1.10$) and neutral trials (M = 535, SD = 107; $t(33) = 9.575$, $p < .001$; $d = 1.08$); and synaesthetes showed overall faster RT compared to non-synaesthetes (M = 542, SD = 103 and M = 620.5, SD = 114, respectively; $d = .73$).

η_p^2 = .39) reflected the presence of a Flanker or congruency effect (CE). Specifically, ER were significantly higher on incongruent (M = 7.48, SD = 9.37) than congruent (M = .43, SD = 1; $t(33) = 4.47$, $p < .001$; $d = 1.06$) and neutral trials (M = .86, SD = 1.79; $t(33) = 4.50$, $p < .001$; $d = .98$) (Bonferroni-adjusted α = .017). No significant differences were found between congruent and neutral trials ($t(33) = 1.65$, $p = .11$). No 'Trial type' x 'Group' interaction emerged to be significant ($F(1.04, 33.2) = 1.57$, $p = .22$) (see Fig. 3A).[6]

In the RT analysis, the presence of CE were confirmed by the significant main effect on 'Trial type' ($F(1.33, 42.5) = 285$, $p < .001$, $\eta_p^2 = .90$). Follow-up pairwise comparisons (Bonferroni-adjusted α = .017) indicated that RT were significantly slower for incongruent trials (M = 397, SD = 40.1) compared to congruent (M = 336, SD = 35.4; $t(33) = 18.6$, $p < .001$; $d = 1.63$) and neutral trials (M = 342, SD = 34.9; $t(33) = 16.4$, $p < .001$; $d = 1.48$). Neutral trials were also found to be significantly slower than congruent trials ($t(33) = 4.08$, $p < .001$; $d = .18$). No differences were found between groups, as indicated by the lack of a significant 'Group' main effect ($F(1, 32) = 2.29$, $p = .14$) or a 'Trial type' x 'Group' interaction ($F(1.33, 42.5) = 1.40$, $p = .25$) (Fig. 3B).

2.2.3.2 Bayesian interference analyses.

The ER analyses provided extreme evidence in favour of the alternative hypothesis for the interaction model (i.e. 'Congruency' x 'Group') in both tasks (CCT: BF_{10} = 886.294; FT: BF_{10} = 1162.53). However, the 'Congruency'-only models were the best performing ones (CCT: BF_{10} = 7760.06; FT: BF_{10} = 2.598.44) and there was inconclusive evidence for the inclusion/exclusion of the interaction terms (CCT: Inclusion BF_{10} = .372; FT: Inclusion BF_{10} = .703). There was also extreme evidence in favour of the alternative hypothesis for the RT interaction model in the FT (BF_{10} = 1.151e+16), but the 'Congruency'-only model was the best

[6] Since the ER variables violated the assumption of normality (as assessed by Shapiro-Wilk), we run an additional non-parametric Friedman χ^2 test for the repeated-measures analysis and Wilcoxon signed-rank Z tests for the pairwise comparisons. The analyses confirmed both the main effect of 'Trial type' ($\chi^2(2) = 40.5$, $p < .001$) and the results observed for the paired-samples comparisons (incongruent vs. congruent trials: $Z(33) = 5$, $p < .001$; incongruent vs. neutral trials: $Z(33) = 0$, $p < .001$; neutral vs. congruent trials: $Z(33) = 51.5$, $p = .10$ – Bonferroni-adjusted α = .017).

performing one (BF_{10} = 2.673e+16) and the inclusion/exclusion of the interaction term presented inconclusive evidence (Inclusion BF_{10} = .444). Regarding the CCT RT analysis, the interaction model (i.e. ‘Congruency’ x ‘Group’) was the best performing one presenting extreme evidence (BF_{10} = 2.673e+16). However, the inclusion/exclusion of the interaction term provided inconclusive evidence (Inclusion BF_{10} = 2.931). The complete analyses can be consulted in Appendix B.

2.2.3.3 Linear mixed model analyses.

For the accuracy rates (AR) analyses, model fit values associated with fixed effects (marginal pseudo-R^2) were very low for both tasks (CCT: marginal R^2 = .06; FT: marginal R^2 = .04). The models (i.e. AR ~ ‘Congruency’ x ‘Group’) for the CCT and the FT were statically significant (both p < .001). ‘Congruency’ significantly predicted AR for both tasks, with higher AR for congruent than incongruent trials (both p < .001). ‘Group’ did not predict AR or interacted with ‘Congruency’ for any of the tasks (all p > .13). Lastly, the random effect of participant was a source of variation in both tasks’ models and these were significantly preferred to analogous no-random effects models (both p < .001). However, conditional model fit values (conditional pseudo-R^2), which consider fixed and random effects, were also low (CCT: conditional R^2 = .11; FT: conditional R^2 = .06).

Regarding the RT analyses, marginal pseudo-R^2 values were low for both tasks (CCT: marginal R^2 = .16; FT: marginal R^2 = .20). The models were statistically significant (both p < .001). ‘Congruency’ significantly predicted RT for both tasks, with slower RT for incongruent compared to congruent trials (both p < .001). ‘Group’ did not predict RT in the CCT task (p = .32), but it had a significant moderating effect on ‘Congruency’. In particular, synaesthetes experienced significantly smaller CE than non-synaesthetes (β = -61.9, SE = 21.8, t(2.84) = -3.53, p = .008), confirming the differences observed in the mixed Analyses of Variance. Non-synaesthetes presented overall faster RT than synaesthetes in the FT, but the factor ‘Group’ did not reach significance (p = .08). ‘Group’ did not have either a moderating effect on

'Congruency' in this task (p = .77). Lastly, the random effect of participant was an important source of variation in both tasks' models, which were significantly preferred to no-random effects models (both $p < .001$). Conditional pseudo-R^2 values were considerable for both tasks (CCT: conditional $R^2 = .52$; FT: conditional $R^2 = .41$).[7] See Appendix C for detailed statistics.

2.2.4 Discussion

To investigate whether synaesthetes have enhanced distractor filtering abilities, we measured different aspects of their attentional skills in two separate conflict tasks. First, we compared synaesthetes and controls' performance on the cross-modal congruency task (CCT), in which a relevant tactile target is always accompanied by an irrelevant visual distractor. In this task, participants have to prioritise one sensory modality over the other and the extent to which visual distractors interfere with the processing of the tactile target can be considered a measure of their intermodal selective attention abilities. In addition, we also measured participants' attentional filtering abilities with the classic Eriksen flanker task (FT). This allowed us to measure their general distractor filtering abilities with a standard task typically used to engage the executive control network of attention, thus contributing to the current debate regarding synaesthetes' general executive skills. Importantly, only -visual synaesthetes were tested in this study (i.e. those with at least one synaesthesia type triggering visual concurrents, such as grapheme-colour synaesthesia or sequence-space synaesthesia) to ensure that their synaesthetic attentional filtering experience matched the sensory modality of the task-irrelevant distractors in our tasks.

The results of the FT revealed no differences between -visual synaesthetes and controls. Robust congruency effects (CE) were observed in both groups with slower responses on incongruent than congruent trials, but the task-irrelevant distractors slowed participants'

[7] Visual inspection of residual plots revealed slight deviations from linearity, homoscedasticity, and normality in all the RT analyses. For that reason, we run the same models as generalised linear mixd models using the Penalised Quasi-Likelihood method (*MASS* package; Venables & Ripley, 2002). The analyses replicated the significance of 'Congruency' as a predictor of RT for both tasks and the moderation effect of 'Group' on congruency for the CCT; no other differences emerged.

performance on incongruent trials in a similar way in both groups. Thus, the sample of synaesthetes and non-synaesthetes selected to take part in this study had comparable general executive control as measured in the FT. This result (or lack thereof) is in line with previous studies that used different classic conflict task (e.g. Stroop task, flanker task) to measure synaesthetes' attentional abilities and failed to report reliable differences with controls in the majority of cases (Mattingley et al., 2001; 2006; Rouw et al., 2013). However, a very different pattern of results emerged in CCT, in which reduced CE were observed in -visual synaesthetes compared to controls. The difference between CE in the two groups was driven by faster response times on incongruent trials in synaesthetes than in non-synaesthetes, while no differences were observed for congruent or neutral trials. This specific pattern of results indicates that -visual synaesthetes were able to select and execute the correct response more quickly than controls when conflicting information coming from different sensory modalities was presented, suggesting that they were better at ignoring the irrelevant visual stimuli of the CCT. This finding is consistent with other studies that investigated synaesthetes' susceptibility to multisensory illusions (sound-induced flash illusion – Neufeld et al., 2012; McGurk illusion – Sinke et al., 2014), showing that synaesthetes experienced fewer multisensory illusions than non-synaesthetes (but see Whittingham et al., 2014 and Brang et al., 2012 for different outcomes).

These results on the FT and CCT were further supported by Bayesian interference and linear mixed model analyses. The Bayesian analyses provided extreme evidence that a model including the interaction between congruency and group best explained the results on the CCT, whereas a model with only the congruency variable suited better the FT. Linear mixed model analyses confirmed the pattern of responses described and revealed that there was an important influence of participants' individual variability, especially in the CCT. Our findings expand the existing literature and provide the first direct evidence that -visual synaesthetes are more efficient than controls at dissociating conflicting information from different sensory modalities in a cross-modal task in which the irrelevant sensory modality matches their synaesthetic concurrents.

Taken together, the results of the present study support the hypothesis that synaesthetes' constant need to ignore their irrelevant synaesthetic percepts is associated with enhanced selective attentional skills. This attentional ability seems to impact synaesthetes' cognitive skills beyond the person's immediate synaesthetic experiences. Crucially, however, this advantage seems to only extend to other types of non-synaesthetic multisensory stimuli, as revealed by the significant advantage observed in a cross-modal congruency task and the lack of effects found in the FT, which measures participants' general executive efficiency. The different pattern of results observed in two seemingly similar conflict tasks might suggest that while the mechanisms responsible for synaesthetic attentional filtering (that is, those underlying the inhibition of irrelevant synaesthetic sensations) are at least partially overlapping with the mechanism engaged during our cross-modal congruency task, they are mostly independent from the mechanisms responsible for the management of other types of perceptual conflict such as those involved in flanker tasks. Indeed, several lines of research seem to suggest that similar mechanisms might be responsible for multisensory perception in the general population and inducer-concurrent associations in synaesthetes (e.g. Sagiv & Ward, 2006). If this is the case, experiencing a synaesthetic concurrent would be equivalent to perceiving a stimulus in an irrelevant sensory modality and synaesthetes might be particularly capable to focus on a certain stimulus modality while ignoring another (intermodal attention). Interestingly, intermodal attention is independent from other attentional mechanisms based on spatial selectivity (e.g. Eimer & Schröger, 1998; Hötting, Rösler, & Röder, 2003). This could explain why no synaesthetic advantages were observed in the FT, in which spatial mechanisms are primarily used to select the target from the distractors (e.g. Fan, Flombaum, McCandliss, Thomas, & Posner, 2003).

One critical question is whether these intermodal filtering abilities can be generalised beyond the modality of synaesthetes' synaesthetic concurrents. As described, in the present study -visual synaesthetes were better than non-synaesthetes at ignoring a task irrelevant visual stimulus that was presented simultaneously to a target in a different sensory modality. Because we assessed -visual synaesthetes is unclear whether synaesthetes who experience non-visual concurrents would show analogous advantages for visual stimuli. Or whether -visual synaesthetes would should benefits for stimuli other than visual (e.g. tactile or auditory). Neuroimaging evidence suggests that the recruitment of parietal areas is shared by different types of synaesthesias (e.g. see Rouw et al., 2011 and Specht, 2012 for reviews). However, specific brain areas are also involved in particular synaesthetic sensations, such as the activation of the colour region V4 in synaesthetic colour experience (e.g. Hubbard et al., 2005; Nunn et al., 2002; Sperling et al., 2006; Steven, 2006; van Leeuwen, Petersson, & Hagoort, 2010). Future studies should directly address this point by assessing the filtering abilities of different types of synaesthetes in different attentional tasks in which the task-irrelevant distractor matches and does not match the sensory modality of their concurrents.

On another note, while all our synaesthetes had visual concurrents, they experienced different types of synaesthesia. We explored the possibility that different types of -visual synaesthetes might show different degrees of filtering abilities. Given that the synaesthetes evaluated in the present study presented either a -colour or a sequence-space synaesthesia (or both) and that previous research observed systematic differences in the visual ability of these two types of synaesthetes (e.g. Ward et al., 2017a), synaesthetes were divided in three groups: colour-synaesthetes (i.e. subjects who experienced synaesthesias producing -colour concurrents; e.g. grapheme-colour synaesthesia), sequence-synaesthetes (i.e. participants who experienced sequence-space synaesthesias; e.g. calendar-forms), and both-synaesthetes (i.e. people who experienced both previous types). An orthogonal 2 x 2 between-subjects factorial design (-colour: yes/no; sequence: yes/no) was used – the no-colour and no-sequence group corresponding to non-synaesthetes (*N* -colour: yes 11/ no 23; sequence: yes 11/ no 23 – i.e. *N* controls = 18, *N* synaesthetes with -colour = 5, *N* = synaesthetes with sequences = 5, *N* = synaesthetes with -colour and sequences = 6). CE measured in the CCT were submitted to a two-way Analysis of Variance. Interestingly, the analysis revealed a significant main effect of sequences ($F(1, 30) = 5.075$, $p = .032$, $\eta_p^2 = .145$), but not of -colour ($F(1, 30) = .52$, $p = .475$) nor a sequences and -colour interaction ($F(1, 30) = .020$, $p = .89$). In particular, the analysis showed that CE were significantly reduced in participants with sequence-space synaesthesia compared to participants without (M = 91.2, SD = 60.7 and M = 157, SD = 66.8, respectively). These findings suggest that a specific subgroup of synaesthetes, namely sequence-space synaesthetes, were the ones with the strongest intermodal attentional filtering advantages. While this appears to be a promising line for future research, it should be highly stressed that this was an exploratory analysis and that the samples were not only small, but also unbalanced (11 sequence-space synaesthetes and 23 without). For this reason, the question of possible differences between different synaesthete types should be further confirmed in future studies with appropriate samples.

In sum, the present study has provided the first evidence that -visual synaesthetes are less affected than non-synaesthetes by the presentation of task-irrelevant visual stimuli when they have to focus on a different sensory modality (as measured in a cross-modal congruency task). This finding suggests that synaesthetes might have enhanced intermodal attentional abilities which allow them to ignore or supress the irrelevant information coming from their synaesthetic concurrents. The present results broaden our understanding of synaesthesia's effects on cognition in a research area which remains largely unexplored.

2.3 Unimodal vs. cross-modal differences when vision is target (Study 2)

2.3.1 Introduction

Study 1 provided the first direct evidence that synaesthetes might be more efficient than non-synaesthetes at dissociating conflicting information from different sensory modalities. In particular, we observed that synaesthetes were better than controls at selecting the location of a tactile target delivered to the index finger (top location) or thumb (bottom location) while additional an irrelevant visual distractor was simultaneously presented at a congruent or incongruent location (Cross-modal Congruency Task; CCT). However, this group difference was not found when participants performed the Flanker Task (FT), a (visual) unimodal conflict paradigm which requires reporting the direction of a central arrow (target) flanked by additional arrows (distractors) pointing to congruent or incongruent directions with respect to the target. Study 1 suggested therefore specific advantages in intermodal attention for synaesthetes.

Moreover, the sensory modality of the distractors in these tasks were matched to the type of filtering that synaesthetes might naturally experience. That is, in Study 1, -visual synaesthetes (i.e. those presenting at least one synaesthesia type involving visual concurrents or triggers such as grapheme-colour or sequence-space synaesthesia) were asked to ignore task-irrelevant visual distractors. Thus, one direct follow-up question is whether the attentional advantages observed are specific to the synaesthetic concurrent sensory modality or expand to different sensory modalities. In other words, we investigated whether -visual synaesthetes present the same attentional efficiency when they must ignore, for instance, tactile or auditory distractors in other cross-modal stimuli combinations.

To address this issue, we compared samples of -visual synaesthetes and matched non-synaesthetes in a modified version of the CCT which reversed the instructions regarding the target-distractor sensory modalities (reversed CCT or rCCT). That is, we asked participants to ignore tactile bursts delivered to the index finger (top) or thumb (bottom) while they had to report the location (top/bottom) of flashes of green light simultaneously presented next to those fingers in the same (index finger, top light or thumb, bottom light; congruent condition) or opposite location (index finger, bottom light or vice-versa; incongruent condition). If -visual synaesthetes intermodal attentional abilities extend to other sensory modalities beyond their concurrents, they should be better than controls at filtering out the task-irrelevant stimuli (touch here) of the rCCT.

In addition, in order to confirm that the synaesthetic advantage is specific to intermodal attention and not to selective attention in general, participants completed a visual unimodal version of the CCT. As noted in Study's 1 *Discussion* (section 2.2.4), intermodal attention is independent from other attentional mechanisms based on spatial selectivity (e.g. Eimer & Schröger, 1998; Hötting et al., 2003), primarily used in the resolution of the FT (Fan et al., 2003). Therefore, the FT might not be a suitable task to assess filtering abilities within modalities in a comparable way to the CCT and this could explain the lack of group differences observed in Study 1. In this newly developed task (visual Unimodal Congruency Task or vUCT), participants were also presented with the CCT set up and were instructed to report the location (top or bottom) of a target stimulus simultaneously delivered with a distractor stimulus that could appear in the same location (top/top or bottom/bottom; congruent condition) or opposite location (top/bottom or bottom/top; incongruent condition). In this case, however, both target and distractors were flashes of light of different colours: green for target stimuli and red for distractors. If -visual synaesthetes are limited to advantages in intermodal attention, they should show no differences compared to non-synaesthetes in the vUCT.

2.3.2 Methods

2.3.2.1 Participants.

A total of 65 subjects participated in this study. All participants reported no known neurological illnesses and normal or corrected-to-normal vision. The study was approved by The University of Edinburgh's Psychology Research Ethics Committee and followed the ethical guidelines laid down in the Helsinki Declaration. Participants were recruited via the University's employment website and convenience sampling, and they received a small monetary compensation (£8/hour). Informed consent was obtained from all participants.

2.3.2.1.1 Synaesthesia screening and classification of participants.

Participants underwent the same synaesthesia screening process specified in Study 1. All participants completed a synaesthesia screening interview (Edinburgh Synaesthesia Screening Assessment; ESSA). Individuals who reported -colour associations (e.g. letters-colours, months-colours, music-colours, etc.) further completed synaesthetic consistency tests for these experiences if a test was available (Synesthesia Battery; Eagleman et al., 2007)[8]. In addition, subjects provided detailed descriptions and drawings for cases of sequence-space synaesthesias and responded to specific projector/associator questions for the different types of synaesthesias reported. See section 2.2.2.1.1 for the complete details.

2.3.2.1.2 Final sample.

Twenty-one -visual synaesthetes and 20 age-matched non-synaesthetes were included in the final sample (demographics are reported in Table 2)[9]. Eleven of these participants had also participated in Study 1 (6 synaesthetes and 5 non-synaesthetes). Two additional synaesthetes were removed from the study for not experiencing any -visual synaesthesia (i.e. synaesthesia types involving -vision as the concurrent modality). As a precautionary measure, we also removed 22 more people who were classified as 'weak' synaesthetes (i.e. individuals who only reported having synaesthetic experiences 'Sometimes' on the ESSA questionnaire and/or just passed synaesthetic consistency tests at the loose threshold, as opposed to 'strong'-synaesthetes, who reported having the experiences 'Always' and passed the tests at the strict threshold; see section 2.2.2.1.1). Twenty-one of these 'weak' participants also experienced at least one synaesthesia type involving vision.

Table 2.
Descriptive, chi-square (χ^2), and t-*statistics of Study 2 groups' demographics.*

Demographics	Synaesthetes	Non-synaesthetes	Statistics
N (male)	19 (2)	14 (6)	$\chi^2(1) = 2.74$, $p = .098$
Age (SD)	23 (2.85)	23 (3.16)	$t(39) = .051$, $p = .96$
Handedness (left, ambidextrous)	19 (1, 0)	20 (0, 1)	$\chi^2(2) = 2$, $p = .37$
Nº of (native) languages* (SD)	1.24 (.44)	1.25 (.44)	$t(39) = .087$, $p = .93$
Level of education** (SD)	2.81 (.75)	2.70 (.86)	$t(39) = .43$, $p = .67$

Note: *N* = Sample size, SD = Standard Deviation.
* Nº of (native) languages: 1 = Monolingual, 2 = Bilingual, 3 = Polylingual.
** Level of education: 1 = High School, 2 = Undergraduate, 3 = Master, 4 = PhD, 5 = Postdoc.

[8] Some participants who reported having -colour experiences that could be assessed with the SB were unable to complete the test due to external time/technical constraints. However, all subjects were thoroughly interrogated about these experiences on the screening interview. Their reports were subsequently analysed and concluded to be consistent with the tested participants' interview accounts as well as with the general properties of synaesthesia.
[9] A power analysis was performed for sample size estimation based on Study 1 results. With error probability α = .05 and target power = .80, the projected total sample needed was *N* = 44 (±22 per group). Thus, our proposed sample of 21 synaesthetes and 20 controls should be adequate for the aims of this study. The analysis was conducted with GPower 3.1 (Faul et al., 2007).

Almost all synaesthetes reported multiple types of synaesthesias (overall average of 10.1 types; range 1-23). Most of them experienced synaesthesias related with -colour (86%) and/or sequence-space synaesthesias (67%). Some also experienced mirror-touch synaesthesia (57%), other -visual synaesthesias such as music-patterns (52%), or ordinal-linguistic personifications (29%). Synaesthetes who completed the synaesthetic consistency tests obtained, on average, a score of .54 points (SD = .23). The descriptions and drawings of sequence-space synaesthesia experiences were analysed in detail and their phenomenology was established consistent with the properties in general and of sequence-space in particular (see section 2.2.2.1.1). All synaesthetes were classified as ‘associators’.

2.3.2.2 Experimental procedure.

As in the previous study, participants completed first the synaesthesia screening tests and questionnaires and then the behavioural tasks; the whole process lasting approximately 60 minutes. The study took place in a dimly lit, sound attenuated room. For the behavioural tasks, participants sat in a comfortable chair and rested their heads in a chinrest to maintain a constant distance from the stimuli displays. Stimuli presentation for both tasks was controlled and responses were recorded via E-Prime 2.0® software and hardware (Serial Response Box 200A®, Psychology Software Tools). Each participant performed two tasks: the reversed Cross-modal Congruency Task (rCCT) and the visual Unimodal Congruency Task (vUCT). The order of the tasks, as well as the stimulus-to-response mappings (see below), were counterbalanced between participants. Before the beginning of each task, participants completed a practice block (12 trials) which was repeated if necessary.

Both tasks used the same set up specifications of the classic Cross-modal Congruency Task (CCT) of Study 1 (see section 2.2.2.2). The only modification was that there were four LED lights attached to the cuboid instead of two; two at the top position and two at the bottom, next to each other (Fig. 4). The rCCT followed the methodological design of the CCT task but the instructions regarding the target and distractor sensory modalities were reversed. That is, participants were asked to attend to the visual stimuli (green flashes) and ignore the tactile bursts. On the other hand, two different visual stimuli were presented in the vUCT. As in the rCCT, the target stimuli (green flashes) consisted of three 50 ms flashes interleaved by two 50 ms offset periods (250 ms total duration). The distractor stimuli were visual flashes with the same timing characteristics but coloured red instead. In both tasks, the target (green flashes) and distractor stimuli (tactile bursts/red flashes) were presented simultaneously, followed by a 1,550 ms empty interval in which responses were collected (total response window of 1,800 ms following stimulus onset), and by a variable inter-trial interval (ITI) randomly selected between 100-500 ms.

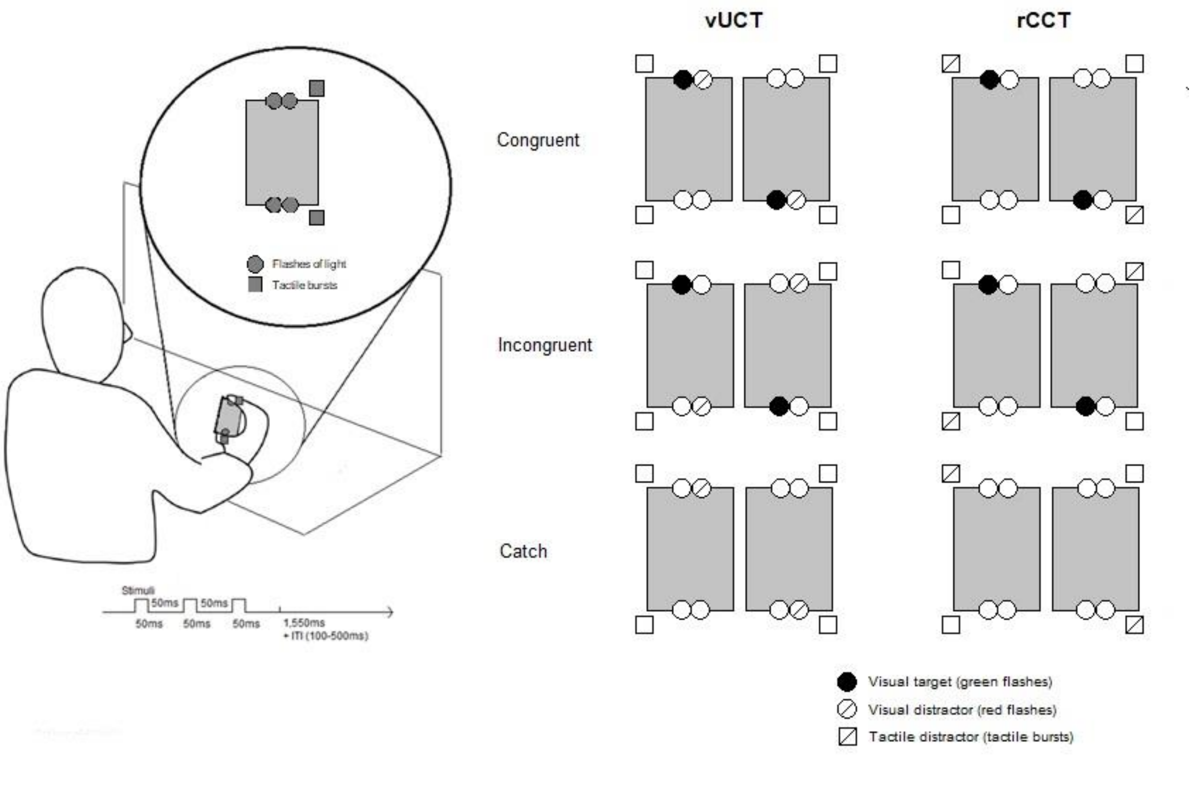


Figure 4. Display, type of trials, and timeline of the visual Unimodal Congruency Task (vUCT) and the reversed Cross-modal Congruency Task (CCT) of Study 2 (adapted from Pavani et al., 2000 and Spence et al., 2004b); ITI = Inter-trial interval.

Congruent and incongruent trials were presented in the tasks. The target and distractors were presented at the same location (top or bottom) on congruent trials, and opposite locations (target top and distractor bottom location, or vice-versa) on incongruent trials. In addition, we included catch trials as a control measure: on those trials only the distractor stimuli were presented (top or bottom location) and thus required a no-response from participants (Fig. 4). We substituted the neutral trials (i.e. target-only trials) from the previous study for this new type of trials given that we did not observe any effects for the former ones in Study 1 (see section 2.2.3.1).

Participants were instructed to perform an elevation discrimination task reporting via pedal press the location (top/bottom) of the targets (green lights) while ignoring the distractors (red lights/tactile bursts). Half of the participants had to press the left pedal with their toes to indicate top location and the right pedal with their heel to indicate bottom location, and the other half followed the opposite mapping. Participants were also instructed to continuously keep their gaze on the fixation point and to answer as rapidly and accurately as possible. They completed five experimental blocks of 80 trials. Within each block, we presented 32 congruent trials, 32 incongruent trials, and 16 catch trials randomly intermixed.

2.3.2.3 Data analyses.

To rule out the possibility of baseline group differences in response criterion, response rates of catch trials were first calculated per task and group and compared with non-parametric independent-samples Mann-Whitney U tests (due to violation of the assumption of normal distribution; as assessed by Shapiro-Wilk). Separate analyses were conducted on error rates (ER) and reaction times (RT). The specifications regarding the inclusion and exclusion of participants and trials in the ER and RT analyses followed the same criteria detailed in Study 1 (see section 2.2.2.3). First, we conducted mixed analyses of variance (ANOVAs) with 'Number of modalities' (unimodal – vUCT task, cross-modal – rCCT task) and 'Congruency' (congruent, incongruent) as within-subjects factors and 'Group' (non-synaesthetes, synaesthetes) as the between-subjects factor. Further pairwise comparisons were carried out as appropriate following significant effects.

To assess the strength of the results observed, we then performed Bayesian interference analyses. In order to simplify the full model, we conducted separate ER and RT Bayesian mixed model ANOVAs for each task using default prior scales and the variable

'Congruency' (congruent, incongruent) as the within-subjects factor and 'Group' (non-synaesthetes, synaesthetes) as the between-subjects factor (see section 2.2.2.3 for details on the defined models and Bayes Factors reported). To measure the influence of participants' individual variability, we also conducted linear mixed model analyses of accuracy rates (AR) and RT as a function of the interaction between 'Congruency' (congruent, incongruent) and 'Group' (non-synaesthetes, synaesthetes), separately for each task. The models included the random effect of participant with random slopes for type of trial (congruent, incongruent) as well. We performed generalised linear mixed models applying the Laplace Approximation method to estimate degrees of freedom and generate *p*-values for the AR analyses, and we run linear mixed models fitting with Maximum Likelihood Estimation and Satterthwaite's method for the RT analyses.

See section 2.2.2.3 for details on the software used to conduct all the different analyses.

2.3.3 Results

2.3.3.1 Mixed Analyses of Variance (ANOVA).

Response rates for catch trials were overall low for both tasks and groups: vUCT (non-synaesthetes M = 1.25, SD = 1.46; synaesthetes: M = 1.19, SD = 1.92) and rCCT (non-synaesthetes M = .75, SD = 1.59; synaesthetes M = .83, SD = 1.74). Furthermore, the analyses showed no differences between groups for neither of the tasks (vUCT: $U(39) = 190$, $p = .58$; rCCT: $U(39) = 202$, $p = .81$).

Outliers and omissions were excluded from the analyses (3.11% of the total trials). Overall, low ER were observed in all groups (non-synaesthetes: M = .59, SD = .67 and synaesthetes: M = .89, SD = 1.66). The analyses revealed a main effect of 'Number of modalities' (i.e. task) ($F(1, 39) = 9.03$, $p = .005$, $\eta_p^2 = .19$), with larger ER on unimodal (vUCT: M = .94, SD = 1.55) than cross-modal trials (rCCT: M = .55, SD = 1.09; $d = .46$) (Fig. 5A). The main effect of 'Congruency' only approached significance ($F(1, 39) = 3.37$, $p = .074$), showing larger ER for incongruent (M = .86, SD = 1.34) compared to congruent trials (M = .63, SD = 1.33) (Fig. 5A). These two factors did not interact (F(1, 39) = .49, p = .49). The factor of interest 'Group' and all the interactions involving it were neither significant (all F(1, 39) < 1.18, all p > .28)[10].

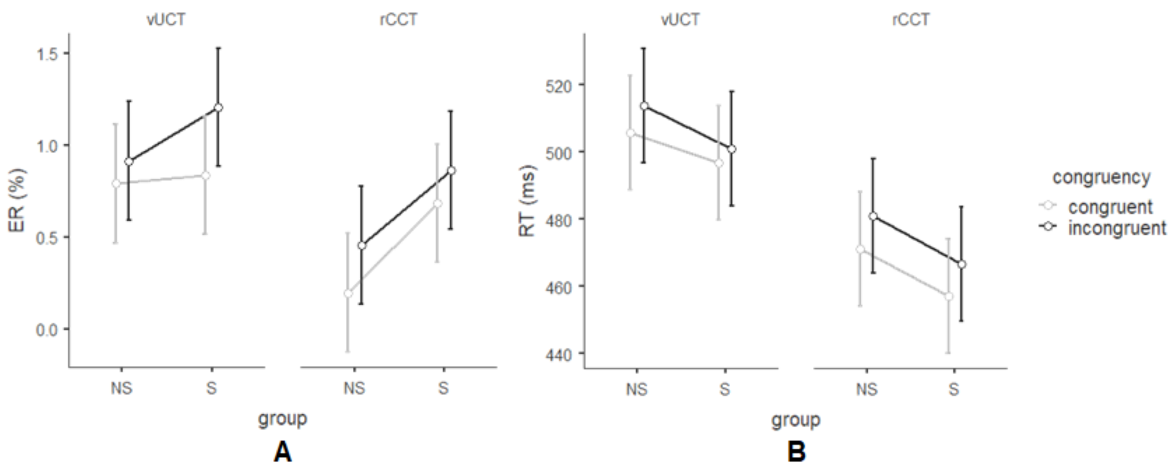


Figure 5. Mean error rates in percentages (ER; Figure **A**) and mean reaction times in milliseconds (RT; Figure **B**) and their corresponding standard error means (error bars) for congruent and incongruent trials, for each task of Study 2 (visual Unimodal Congruency Task – vUCT; reversed Cross-modal Congruency Task – rCCT) and group (non-synaesthetes – NS; synaesthetes – S). The analyses only revealed a main effect of 'Congruency' in the RT analysis (p < .001), evidencing typical congruency effects (i.e. slower RT for incongruent than congruent trials); but it only approached significance in the ER analysis (p = .074).

The RT analyses showed as well a main effect of 'Number of modalities' (F(1, 39) = 34.6, p < .001, η_p^2 = .47), with slower RT for unimodal (vUCT: M = 504, SD = 77.6) compared to cross-modal trials (rCCT: M = 469, SD = 74.3; d = .93). In addition, the factor of 'Congruency' reached significance in these analyses (F(1, 39) = 31, p < .001, η_p^2 = .44) (Fig. 5B). As expected, incongruent trials were significantly slower (M = 491, SD = 75.3) than congruent trials (M = 483, SD = 72.1; d = .87) (Fig. 5B). The interaction between these two factors was not significant (F(1, 39) = 1.87, p = .18). There was not a main effect of 'Group' or interactions concerning this factor (all F(1, 39) < .48, all p > .50; Fig. 5B).[11]

[10] Since the ER variables violated the assumption of normality (as assessed by Shapiro-Wilk), we run additional non-parametric repeated-measures Friedman χ^2 tests. The analyses showed a trend for the main effect of 'Number of modalities' (χ^2(1) = 3.46, p = .006) and no effect for 'Congruency' (χ^2(1) = .47, p = .49).

[11] Since some of the participants of Study 2 had also participated in Study 1 (see section 2.3.2.1.2), we conducted additional analyses to disregard the possibility that the performances of these subjects and of the new participants presented any differences. To that aim, we run independent samples t-tests comparing congruency effects (i.e. differences in milliseconds/error rates between incongruent and congruent trials) between those subjects who had participated in Study 1 and those who had not, separately for each group, task, and measurement. The analyses

2.3.3.2 Bayesian interference analyses.

The ER analyses provided substantial to strong evidence in favour of the null hypothesis regarding the interaction model (i.e. ‘Congruency’ x ‘Group) in both tasks (vUCT: BF_{10} = .07; rCCT: BF_{10} = .186). The ‘Congruency’-only models were the best performing ones in the RT analyses with substantial evidence for the vUCT task (BF_{10} = 8.86) and extreme evidence for the rCCT task (BF_{10} = 2852.57). There was also moderate evidence for the alternative hypothesis with respect to the interaction model in the vUCT (BF_{10} = 3.208) and extreme evidence in the rCCT (BF_{10} = 768.013). However, the inclusion/exclusion of the interaction term presented inconclusive evidence (vUCT: Inclusion BF_{10} = .406; rCCT: BF_{10} = .327). The complete analyses can be consulted in Appendix B.

2.3.3.3 Linear mixed model analyses.

For the accuracy rates (AR) analyses, model fit values associated with fixed effects (marginal pseudo-R^2) were very low for both tasks (vUCT: marginal R^2 = .002; rCCT: marginal R^2 = .001). The models (i.e. AR ~ ‘Congruency’ x ‘Group’) for the vUCT and the rCCT were statically significant (both p < .001). ‘Congruency’ did not predict AR in the vUCT (p = .24) but showed a predictive trend in the rCCT (p = .06), with higher AR for congruent than incongruent trials. ‘Group’ did not predict AR or interacted with ‘Congruency’ for any of the tasks (all p > .24). Lastly, the random effect of participant was a source of variation in both tasks’ models and were significantly preferred to analogous no-random effects models (both p < .001). However, conditional model fit values (conditional pseudo-R^2), which consider fixed and random effects, were also very low (vUCT: conditional R^2 = .02; rCCT: conditional R^2 = .01).

Regarding the RT analyses, marginal pseudo-R^2 values were very low for both tasks (vUCT: marginal R^2 = .002; rCCT: marginal R^2 = .005). The models were statistically significant (both p < .001). ‘Congruency’ significantly predicted RT for both tasks, with slower RT for

revealed no differences in RT or ER for any of the groups or tasks (synaesthetes: all $t(19)$ < 1.24, all p > .23; non-synaesthetes: all $t(18)$ < 1.15, all p > .27).

incongruent compared to congruent trials (both $p < .004$). 'Group' did not predict RT or interacted with 'Congruency' for any of the tasks (all $p > .44$). Lastly, the random effect of participant was a source of variation in both tasks' models, which were significantly preferred to no-random effects models (both $p < .001$). Conditional pseudo-R^2 values were moderate for both tasks (CCT: conditional $R^2 = .34$; FT: conditional $R^2 = .38$). [12] See Appendix C for the detailed statistics.

2.3.4 Discussion

The present study followed up on Study's 1 findings that synaesthetes might be better than non-synaesthetes at filtering irrelevant and potentially conflicting information from different sensory modalities when the modality of the irrelevant stimuli coincides with the modality of their synaesthetic concurrent. For instance, synaesthetes experiencing visual associations such as colours for letters have an advantage at ignoring visual irrelevant stimuli. In particular, Study 2 addressed the question whether synaesthetes' attentional efficiency extends to other sensory modalities that do not match their synaesthetic concurrents (e.g. whether letters-colours synaesthetes are also better at filtering irrelevant tactile or auditory stimuli). To investigate this, a group of synaesthetes who experienced synaesthetic -visual concurrents were compared to non-synaesthetes in a cross-modal conflict task in which they had to attend to visual targets while ignoring simultaneous tactile distractors (reversed Cross-Modal Congruency Task or rCCT). The degree of interference of the distractors in relation to the processing of the targets can be regarded as a measure of subjects' selective attention abilities. In addition, participants completed a visual unimodal version of the same task (visual Unimodal Congruency Task or vUCT) in which distinct visual targets and distractors were presented. The task aim was to confirm the absence of synaesthetic advantages in filtering stimuli within the same sensory modality.

[12] Visual inspection of residual plots revealed slight deviations from linearity, homoscedasticity, and normality in all the RT analyses. For that reason, we run the same models as generalised linear mixd models using the Penalised Quasi-Likelihood method (*MASS* package; Venables & Ripley, 2002). The analyses replicated the significance of 'Congruency' as a predictor of RT for both tasks; no other differences emerged.

In both tasks, irrelevant stimuli caused strong interference effects, as evidenced by congruency effects (CE) or slower reaction times / higher error rates when the target and the distractor stimuli were presented in opposite locations (incongruent condition) compared to when they were presented in the same location (congruent condition). However, results did not reveal any group differences in CE size. Thus, synaesthetes and non-synaesthetes processed the task-irrelevant distractors in a comparable way. Bayesian interference analyses determined that models with only the congruency variable explained better the results on these tasks. In addition, linear mixed model analyses confirmed the pattern of responses described and revealed that there was an important influence of participants' individual variability in both tasks.

On the one hand, the unimodal task (vUCT) was designed to eliminate the spatial selectivity requirements present in the Flanker Task assessed in Study 1 (see section 2.2.4) and thus make it more comparable to the CCT, in which synaesthetes were observed to be better than controls at filtering visual distractors paired with tactile targets (i.e. intermodal attention). Therefore, the null results in the vUCT do not contribute to the hypothesis that synaesthetes might experience the same attentional advantages when they are required to dissociate conflicting information within the same modality (i.e. selective attention). On the other hand, synaesthetes did not differ either from non-synaesthetes when it came to process tactile distractors in the rCCT. Both in Studies 1 and 2 synaesthetes experienced -visual concurrents (i.e. at least one synaesthesia type involving vision as the concurrent sensory modality such as letters-colours synaesthetes). This is important because it means that in Study's 1 classic CCT, the sensory modalities of the task-irrelevant stimuli and synaesthetes' concurrents matched, whereas this was not the case for Study's 2 rCCT. Thus, the lack of group differences in the rCCT task cannot support the hypothesis that synaesthetes' enhanced intermodal attention abilities might extend to sensory modalities beyond those that match their synaesthetic concurrents.

Despite the fact that the vUCT and rCCT tasks produced much smaller CE compared to the classic CCT (average of 11.5 and 11.6 ms compared to 136 ms, respectively), the effects were robust and significantly strong. As mentioned, the linear mixed model analyses highlighted the influence of participants' individual variability, but this was the case as well in Study's 1 tasks. In addition, subjects followed the same screening procedure in both studies. Therefore, it is unlikely that the lack of group differences of the present study is due to issues with its design or sample. It is also important to note that Bayesian interference analyses provided substantial to extreme evidence in favour of explanation models that did not include group as a factor. Therefore, although Study 2 results seem to reinforce the notion that -visual synaesthetes might be specifically better at filtering out irrelevant information from different sensory modalities when the irrelevant modality matches their synaesthetic concurrents, no definitive conclusions can be drawn from the present results.

2.4 Visuo-tactile vs. audio-visual differences: Vision as target vs. vision as distractor (Study 3)

2.4.1 Introduction

-Visual synaesthetes, compared to controls, seem to have specific advantages at filtering conflicting information from different sensory modalities when the irrelevant modality matches their synaesthetic concurrent modality (i.e. for instance, synaesthetes grapheme-colours or sequence-space synaesthetes are more efficient at ignoring visual irrelevant stimuli). In Study 1 we observed that these synaesthetes were better than non-synaesthetes at ignoring the visual distractors of a visuo-tactile task (Cross-modal Congruency Task or CCT). One important question is whether this synaesthetic attentional advantage is consistent across different sensory modalities combinations. In other words, whether the modality of the target is a relevant variable and the same differences in attention can be observed when synaesthetes must ignore visual distractors presented, for instance, with auditory targets.

A second aim of this study was to explore the possibility that different types of synaesthetes might show different attentional abilities. Both Studies 1 and 2 prominently included two -visual synaesthetes subgroups: participants who experienced -colour synaesthesias (e.g. letters-colours, months-colours, etc.) and participants who experienced sequence-space synaesthesias (e.g. calendar-forms, number-lines, etc.). Interestingly, exploratory analyses in Study 1 suggested particularly strong intermodal attention abilities for sequence-synaesthetes compared to colour-synaesthetes (see section 2.2.3). However, the subsamples analysed there were too small to draw any firm conclusions. Several studies seem to indicate though that sequence-space synaesthetes might have indeed specific cognitive characteristics compared to non-synaesthetes and other types of synaesthetes. Sequence-synaesthetes have showed advantages in tasks involving memory (Brang et al., 2010; Lunke & Meier, 2018; Simner et al., 2009b), time manipulation (Mann et al., 2009), mental rotation and visual imagery (Brang, Miller, McQuire, Ramachandran, & Coulson, 2013a; Lunke & Meier, 2018; Simner et al., 2009b), spatial processing (Hale, Thompson, Morgan, Cappelletti, & Cohen-Kadosh, 2014), or visual perception (Ward et al., 2017a). In addition, several studies have reported higher rates of (self-reported) visual imagery for this subgroup of synaesthetes (Havlik et al., 2015; Price, 2009; Spiller & Jansari, 2008; Spiller et al., 2015; Ward et al., 2018a). However, experiencing this type of synaesthesia also seems to have some costs. Sequence-synaesthetes were more affected than control groups when they were assessed on cognitive flexibility related to number and time manipulations (Ward, Sagiv, & Butterworth, 2009) or when they were asked to make temporal judgements about the order of the months and numbers incompatible with their spatial forms (Price & Mentzoni, 2008; Gertner, Henik, & Cohen-Kadosh, 2009; Hubbard, Ranzini, Piazza, & Dehaene, 2009; Smilek, Callejas, Dixon, & Merikle, 2007a).

The present study addressed these questions by comparing matched groups of colour-synaesthetes (i.e. subjects who experienced synaesthesias involving -colour as the concurrent; e.g. grapheme-colour synaesthesia), sequence-synaesthetes (i.e. subjects experienced sequence-space synaesthesia forms; e.g. calendar-forms), and non-synaesthetes in different versions of the original CCT. On the one hand, tasks presented either visuo-tactile or audio-visual stimuli. On the other hand, the sensory modality of the distractors could be concurrent-related (i.e. visual) or -unrelated (i.e. tactile or auditory). Each participant performed the four resulting tasks: visuo-tactile concurrent-unrelated (VTCU; vision-target, touch-distractor), audio-visual concurrent-unrelated (AVCU; vision-target, audition-distractor), visuo-tactile concurrent-related (VTCR; touch-target, vision-distractor), and audio-visual concurrent-related (AVCR; audition-target, vision-distractor). We included the visuo-tactile concurrent-related and -unrelated tasks (i.e. VTCR and VTCU, equivalent to the CCT and rCCT, respectively) for replication purposes and in order to be able to compare directly the different questions evaluated in this study with a new sample of synaesthetes. Similarly, we decided to add the audio-visual concurrent-unrelated (AVCU) version to ratify the absence of attentional advantages when the distractor and the synaesthetic concurrent modalities do not match in a new combination of sensory modalities. If -visual synaesthetes' intermodal abilities are consistent across different sensory modality combinations, they should be better than non-synaesthetes at filtering the irrelevant stimuli of the two concurrent-related tasks (VTCR and AVCR). In line with previous investigations, we also hypothesise that this advantage should not be observed for the concurrent-unrelated tasks (VTCU and AVCU). Finally, we expect to find stronger advantages for sequence-synaesthetes compared to colour-synaesthetes.

2.4.2 Methods

2.4.2.1 Participants.

This study was part of a larger data collection project carried out in collaboration with the University of Sussex (Brighton, UK). The project involved three different sessions in which 359 total subjects participated in all or part of these sessions. All participants reported no known neurological illnesses and normal or corrected-to-normal vision. The project was approved by The University of Edinburgh´s Psychology Research Ethics Committee and the University of Sussex's Sciences & Technology Cross-Schools Research Ethics Committee, and followed the ethical guidelines laid down in the Helsinki Declaration. Participants were recruited via The

University of Edinburgh's employment website and they received a small monetary compensation (£8/hour). Informed consent was obtained from all participants.

In the first session, participants had to complete on-line a questionnaire version of the synaesthesia screening interview used in Studies 1 and 2 and that we revised here (see the following section and Chapter IV). In addition, they completed several personality tests (see Chapter III). The second session was conducted in the Cross-Modal Lab of The University of Edinburgh and consisted of completing the revised synaesthesia screening interview and a series of questionnaires and tests aimed at assessing synaesthetic experiences in more detail (see the following section). Finally, the third session, carried out in the Lab as well, consisted in completing the behavioural tasks (see section 2.4.2.2). The present study only includes those participants with complete data for both the second (lab screening) and third (behavioural tasks) sessions, which amounted to a total of 122 participants.

It should be noted that the participants' recruitment advert was not aimed at synaesthetes. Special care was taken not to include the word 'synaesthesia' or its definition in the study ad description or in the instructions of the synaesthesia screening questionnaire completed on-line. Interested subjects were invited to participate in a "perception experiment" and they were told that they would have to complete questionnaires that would "explore their perceptual experiences" in addition to other questionnaires and computerised tests. Moreover, the researcher did not examine the answers to the first session's on-line questionnaires until all the data collection process was finished; everyone that successfully completed the on-line screening questionnaire was contacted to schedule the next session in the lab.

2.4.2.1.1 Synaesthesia screening and classification of participants.

In this study we updated the synaesthesia screening protocol. First, participants were interviewed with a revised version of the Edinburgh Synaesthesia Screening Assessment (ESSA) used in Studies 1 and 2 (see section 2.2.2.1.1). As its former version, the new screening tool explored different types of synaesthetic experiences reported in the literature.

However, instead of inquiring about different aspects of these experiences (i.e. frequency, constancy, location, and stability), the person was now asked to rate how much each type of synaesthesia applied to him/her with a 5-point Likert scale ranging from 'Not at all' to 'Completely'. Complete details about the revised ESSA, which participants also completed as a self-report questionnaire in the first on-line session, can be found in Chapter IV and consulted in Appendix D.

Following next, participants completed synaesthetic consistency tests for -colour and sequence-space synaesthesias. For -colour experiences, subjects completed the Synesthesia Battery (SB) used in previous studies (see section 2.2.2.1.1) or the Sussex's Multisense Consistency Test (MCT for short; MULTISENSE Research Project, 2019)[13]. Like the SB, the MCT is a standardised tool which measures the person´s consistency at choosing the same colour for each synaesthetic trigger presented several times (e.g. the different digits 0 to 9) and offers tests for letters, numbers, weekdays, and months. In the SB, scores below 1.0 (or 1.43 in the looser threshold; see section 2.2.2.1.1) are considered to indicate presence of synaesthesia. The MCT provides comparable SB scores: an average score of ≥85% in the MCT is equivalent to a score of <1.0 in the SB, and an average score of ≥75% to a <1.43.

All participants did the letters-colours test as a control measure and were instructed to "choose the colour that you believe that fits each letter best". People that claimed to have -colour associations for letters were given the option to select 'no colour' if they did not associate a colour for a specific letter, whereas the rest of subjects were told to ignore that option. In addition to the letters-colours test, those subjects that expressed having synaesthetic experiences for any of the other -colour tests available, completed them as well. If a person took more than one test, the minimum average score obtained in any of them (SB scoring system) was taken as the reference score for classification purposes.

[13] The use of different consistency tests for -colour experiences was forced by the closure of the website hosting the Synesthesia Battery (SB) midway the data collection process.

Regarding the consistency test for spatial sequences, participants completed the Sussex's Sequence-Spatial Synaesthesia Diagnostic Test (SDT for short; Ward et al., 2018a). In this test, participants are asked to place each day of the week, month, and digits 0-9 on a blank computer screen over repeated trials (three per item, presented in a randomised order). In a similar fashion to the -colour consistency tests, the different locations (measured in x and y coordinates) given for each item are computed into a three-point measurement (i.e. area) and then averaged across all items. The smaller the average area, the stronger the synaesthetic association is considered (as it reflects more consistency at giving the same locations for each particular item).

To be considered a sequence-space synaesthete, subjects need to obtain an average area below .203 and a standard deviation above .1 for either the x or the y coordinate (as a low number for both values would indicate that the person has tended to click in the same place for all items). Ward et al. (2018a) also proposed an alternative, looser cut-off of <.300 for the average area score, producing only modest effects on the prevalence estimates. In addition to the average area criteria, participants must pass a short questionnaire aimed at quantifying the subjective experience of sequence-space synaesthesia (see Ward et al., 2017a for further details). Total scores range between 9 and 45 points and low scores reflect synaesthesia-alike answers; a total score of or below 25 points[14] being indicative of the presence of synaesthesia. Thus, in our study, sequence-synaesthetes were those who

[14] As noted here, genuine sequence-synaesthetes are identified as those who score, inter alia, within a required range on a self-report questionnaire (< 25). This test for sequence-space synaesthesia was taken from Ward et al. (2017), who validated the test on a group of known sequence-synaesthetes. In that version of the test, the required score in the questionnaire was < 19. However, a minor change to the wording in our own version of the test required us to re-set the threshold for synaesthesia here to < 25. The change in wording we introduced (we clarified that spatial forms did not come from something seen in a book; e.g. "Before doing this experiment, I had always thought about NUMBERS as existing in a particular sequence (i.e. I've always had my own number-space, and it's not just something I've seen in a book […])") appeared to result in more hesitancy in fully agreeing with our synaesthete-like statements. As a result, our questionnaire pilot here (M = 30.5 and SD = 8.71) was higher than reported in Ward et al. (Ward et al., 2018a; M = 24.1, SD = 7.40). We therefore standardised our threshold against the threshold validated by Ward et al. (2018a). In this latter questionnaire, the threshold fell at a Z = .69. Applying the same standardised threshold to our own data produces a revised threshold of 24.6, which rounds to 25 given that all possible scores are whole integers.

satisfied these two independent requirements: scoring ≤25 on the questionnaire and ≤.203 (or ≤.300 according to the loose criterium) in the consistency test.

Following the same rationale as in the -colour consistency test for letters, all participants completed the SDT as a control measure. Regardless of whether people reported or not to experience spatial forms, all subjects received the same instructions: “Think how [the different numbers, days of the week, and months] may be arranged spatially on the 2D computer screen”. As in Studies 1 and 2, participants claiming to have sequence synaesthesias were asked to provide detailed descriptions and drawings of these experiences. In this study, subjects who did not experience them were also told to: “Think how the months of the year might be arranged spatially and draw them”. In addition, projector/associator questions were asked for sequence experiences and the rest of synaesthesia types declared following the previous studies procedure (see section 2.2.2.1.1). Lastly, participants also completed the Edinburgh Handedness Inventory (Oldfield, 1971).

Participants were classified as non-synaesthetes if they failed both consistency tests (according to the strict and loose criteria) *and* did not report any type of synaesthetic experience in the synaesthesia screening interview (ESSA). To be classified as synaesthetes, subjects had to pass the -colour and/or sequence-space consistency tests (according to the strict criteria). For the rare cases of participants passing the consistency tests but not reporting any degree of synaesthetic experience on the interview, we decided to classify them as synaesthetes anyway because consistency measures are already robust against non-synaesthetic answer patterns (and thus we considered that failure to report these experiences responded to lack of awareness). People failing the consistency tests at the strict criteria but passing either or both of them at the loose threshold were categorised into a different group called ‘weak-synaesthetes’. Lastly, participants failing the consistency tests at the strict and the loose criteria *but* claiming to have other types of synaesthesias that could not be assessed with consistency tests, were categorised into a fourth group called ‘other-synaesthetes’. This included other potential colour- or -visual synaesthetes (e.g. pain-colours or music-patterns) as well as people with alternative types of synaesthesias (e.g. mirror-touch).

2.4.2.1.2 Final sample.

A total of 122 people completed the lab screening and behavioural sessions of the data collection project (see section 2.4.2.1). Twenty-eight participants were classified as non-synaesthetes and 55 as synaesthetes: 22 colour-synaesthetes, 21 sequence-synaesthetes, and 12 who presented both types[15]. In addition, 15 more subjects were classified as weak-synaesthetes and another 24 as other-synaesthetes. These individuals were excluded from the analyses because they did not fit our sample inclusion criteria. Synaesthetes with both types of synaesthesias were also removed in order to be able to properly examine whether different types of synaesthesias might present specific attentional abilities. Table 3 shows the main descriptive statistics for the Study 3 final sample and Table 4 a summary of the average consistency scores obtained per group.

Table 3.
Descriptive, chi-square (χ^2), and t-*statistics of Study 3 groups' demographics.*

Demographics	Colour-synaesthetes	Sequence-synaesthetes	Non-synaesthetes	Statistics
N (male)	16 (6)	20 (1)	17 (11)	$\chi^2(2) = 7.62$, $p = .022$
Age (SD)	21.5 (2.28)	21.4 (2.23)	22.5 (3.04)	$F(2, 64) = 1.07$, $p = .35$
Handedness (left, ambidextrous)	18 (3, 1)	19 (0, 2)	26 (2, 0)	$\chi^2(2) = 5.60$, $p = .23$
Nº of (native) languages* (SD)	1.36 (.58)	1.24 (.44)	1.14 (.36)	$F(2, 64) = 1.97$, $p = .14$
Level of education** (SD)	2.50 (.51)	2.38 (.50)	2.39 (.57)	$F(2, 64) = .19$, $p = .83$

Note: *N* = Sample size, SD = Standard Deviation.
* Nº of (native) languages: 1 = Monolingual, 2 = Bilingual, 3 = Polylingual.
** Level of education: 1 = High School, 2 = Undergraduate, 3 = Master, 4 = PhD, 5 = Postdoc.

[15] A power analysis was performed for sample size estimation based on Study 1 results. With error probability α = .05 and target power = .80, the projected total sample needed was *N* = 54 (±18 per group). Thus, our proposed sample of 22 colour-synaesthetes, 21 sequence-synaesthetes, and 28 controls should be adequate for the aims of this study. The analysis was conducted with GPower 3.1 (Faul et al., 2007).

Table 4.
-Colour and sequence synaesthetic consistency scores of Study 3 participants by group.

Synaesthetic consistency measures	Colour-synaesthetes	Sequence-synaesthetes	Non-synaesthetes
-Colour SB/MCT consistency score (SD)*	.71 (.18)	1.70 (.52)	2.31 (.43)
Sequence SDT consistency score (SD)	.84 (.99)	.09 (.05)	.59 (.90)
Sequence SDT questionnaire score (SD)	29.4 (7.85)	18.6 (3.70)	37.9 (5.54)

Note: SD = Standard Deviation, SB = Synesthesia Battery (synaesthetic threshold < 1.0), MCT = Multisense Consistency Test (synaesthetic threshold ≥ 85%), SDT = Sussex's Sequence-Spatial Synaesthesia Diagnostic Test = SDT (consistency threshold < .203; questionnaire threshold < 25).
* Scores of participants who completed the MCT were transformed to SB scores for homogenisation purposes.

Almost all synaesthetes reported experiencing multiple types of synaesthesias, with an overall average of 12.6 types (range: 2-69)[16]. Besides -colour and/or sequence-space synaesthesias, 45.4% of the synaesthetes claimed to experience other -visual synaesthesias not involving colour (e.g. music-patterns or ticker-tape synaesthesia) and 44.2% of them mirror-type synaesthesias such as mirror-pain synaesthesia. Other types of synaesthesias registered were ordinal-linguistic personifications (27.9%) and experiences related with -touch (23.3%), -pain (11.6%), -smell (11.6%), -sound (11.6%), and -taste (9.3%).

The descriptions and drawings of sequence-space synaesthesia experiences were analysed in detail and their phenomenology was established consistent with the properties of synaesthesia in general and of sequence-space in particular (see section 2.2.2.1.1). The corresponding control drawings (see section 2.4.2.1.1) made by participants who did not experience this type of synaesthesia were also inspected and deemed not to fit the previous criteria. All synaesthetes were classified as 'associators'.

2.4.2.2 Experimental procedure.

The session took place in a dimly lit, sound attenuated room. Participants sat in a comfortable chair and rested their heads in a chinrest to maintain a constant distance from the stimuli

[16] Without considering an extreme case synaesthete reporting 69 different synaesthetic experiences, the overall average of number of synaesthesia types was 11.2 with a range of 2-33.

displays. Stimuli presentation was controlled and responses were recorded via E-Prime 2.0® software and hardware (Serial Response Box 200A®, Psychology Software Tools).

Each participant performed four conflict tasks that resulted from different combinations of sensory stimuli (visuo-tactile or audio-visual) and different specifications regarding the target and distractor modalities (vision-target or vision-distractor). For the visuo-tactile and audio-visual concurrent-related versions (VTCR and AVCR, respectively), the irrelevant stimuli matched the synaesthetic concurrent experiences of our participants. Therefore, they had to attend to either tactile or auditory targets while ignoring visual distractors. On the other hand, in the visuo-tactile and audio-visual concurrent-unrelated versions (VTCU and AVCU, respectively) the target-sensory modalities were reversed – i.e. they had to attend to visual targets while ignoring tactile or auditory distractors. The concurrent-related and -unrelated versions of each sensory pairing presented the same stimuli; the only difference were the instructions given to the participant (i.e. whether they had to attend to one modality or to the other). The order of the tasks, as well as the stimulus-to-response mapping (see below), were counterbalanced between participants. Before the beginning of each task, participants completed a practice block (40 trials) which was repeated if necessary. The study lasted approximately 60 minutes.

The visuo-tactile tasks used the same set up specifications of the Cross-modal Congruency Task (CCT) of Study 1 (see section 2.2.2.2). Participants were instructed to perform an elevation discrimination task reporting via pedal press the location (top/bottom) of the targets (VTCU: green flashes or VTCR: tactile bursts) while ignoring the distractors (VTCU: tactile bursts or VTCR: green flashes). Thus, the VTCR corresponded to the classic CCT whilst the VTCU was a slightly modified version of the reversed CCT (rCCT) assessed in Study 2 (i.e. two LED lights – top and bottom – instead of the four – two top, two bottom – used in Study 2; see section 2.3.2.2). Half of the participants had to press the left pedal with their toes to indicate top location and the right pedal with their heel to indicate bottom location, and the other half followed the opposite mapping.

For the audio-visual tasks, the stimuli were presented using a 17-inches Dell® UltraSharp 1708FP computer screen and two Zalman® ZM-S300 speakers. The computer screen was positioned in front of the participants (50 cm from the table edge where the chinrest was attached) and aligned with their body midline. The left and right speakers were respectively placed next to the left and the right sides of the computer screen and elevated at the same height (15 cm from the base). On each trial, an auditory target/distractor was presented either to the left or the right speaker and consisted of three 60 ms onset periods of monaural tones at a sampling rate of 44.1 kHz and 60 dB(A), interleaved by two 60 ms offset period (300 ms total duration)[17]. The simultaneous visual target/distractor consisted of a white asterisk (*) symbol (subtended visual angle of 2.76°) that appeared on a black background, 2.29° from the left or right sides of the computer screen, centred. A central white cross (subtended visual angle of 1.15°) was present throughout the duration of each task and served as fixation point. Each trial started with the presentation of the stimuli (300 ms), followed by a 1,500 ms empty interval in which responses were collected (total response window of 1,800 ms following stimulus onset), and a variable inter-trial interval (ITI) randomly selected between 100 and 500 ms. Participants were instructed to report via pedal press the location (left/right) of the targets (AVCU: tones and AVCR: asterisks) while ignoring the distractors (AVCU: asterisks and AVCR: tones) (Fig. 6). Given the different pedal positions needed for the visuo-tactile tasks (see above), half of the participants had to press the left pedal pressing with their toes and the right pedal pressing with their heel and vice-versa for the other half.

[17] We tried to keep the designs of the visuo-tactile and the audio-visual tasks as comparable as possible, but some slight modifications in the specifications had to be applied due to technical constraints.

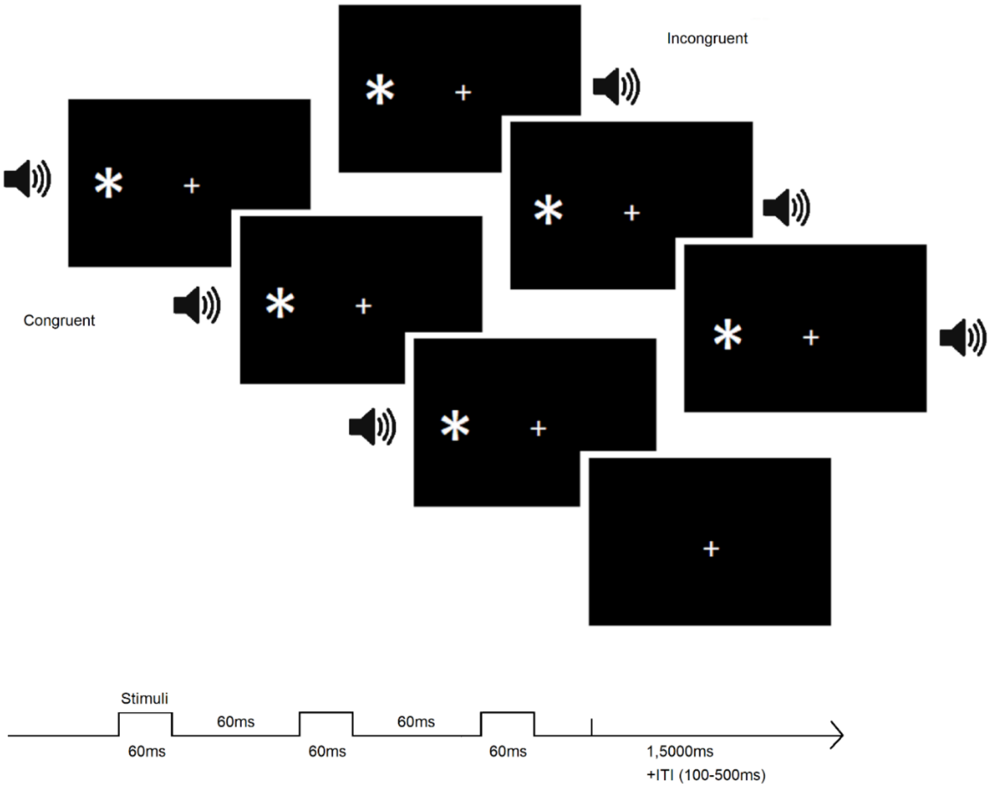


Figure 6. Display, type of trials, and timeline of the audio-visual tasks of Study 3 (Audio-Visual Concurrent-Unrelated – AVCU; Audio-Visual Concurrent-Related – AVCR); ITI = Inter-trial interval.

In all tasks, two different types of trials were presented: congruent and incongruent. The target and the distractor stimuli were simultaneously presented from the same location (top or bottom or left or right) on congruent trials, and from opposite locations (target stimulus top or left and distractor stimulus bottom or right, or vice-versa) on incongruent trials. We removed the catch trials (i.e. distractor-only trials) from the previous study given that all participants showed a low response rate and that there were no differences between groups (see section 2.3.3.1). Participants were instructed to answer as rapidly and accurately as possible and to keep their gaze on the fixation point at all times. For each task, participants completed two experimental blocks of 80 trials. Within each block, congruent and incongruent trials were equally likely (40 trials per type) and randomly intermixed.

2.4.2.3 Data analyses.

Separate analyses were conducted on error rates (ER) and reaction times (RT). The specifications regarding the inclusion and exclusion of participants and trials in the ER and RT analyses followed the same criteria detailed in Study 1 (see section 2.2.2.3). First, we conducted mixed analyses of variance (ANOVAs) with the within-subjects factors of ‘Target-distractor’ (vision-target, vision-distractor), ‘Sensory modalities’ (visuo-tactile, audio-visual), and ‘Congruency’ (congruent, incongruent); and the between-subjects factor of ‘Group’ (non-synaesthetes, colour-synaesthetes, and sequence-synaesthetes)[18]. Further paired- and independent-samples *t*-tests were carried out as appropriate following significant interactions. To simplify complex interactions involving the factor ‘Congruency’, the relevant follow-up analyses were carried out on the congruency effects or CE (i.e. differences in RT or ER between incongruent and congruent trials). In case of violation of the assumption of normal distribution of the dependent variables (as assessed by Shapiro-Wilk), we used alternative, non-parametric tests: Wilcoxon signed-rank Z tests for two-group paired-samples comparisons and Mann-Whitney U tests for two-group independent-samples comparisons. In addition, Bonferroni corrections for multiple comparisons were applied adjusting the alpha threshold accordingly when needed.

To assess the strength of the results observed, we then performed Bayesian interference analyses. With this purpose in mind and given the complexity of the full model, we decided to conduct separate RT and ER Bayesian mixed model ANOVAs for each task. We used default prior scales and the variable ‘Congruency’ (congruent, incongruent) as the within-subjects factor and ‘Group’ (non-synaesthetes, colour-synaesthetes, and sequence-synaesthetes) as the between-subjects factor (see section 2.2.2.3 for details on the defined models and Bayes Factors reported). In order to measure the influence of participants’ individual variability, we also conducted linear mixed model analyses of accuracy rates (AR) and RT as a function of the interaction between ‘Congruency’ (congruent, incongruent) and ‘Group’ (non-synaesthetes, colour-synaesthetes, sequence-synaesthetes), separately for each task. The models also included the random effect of participant with random slopes for

[18] To disregard any possible influences of gender (see Table 3), we run the same mixed ANOVAs for ER and RT adding this variable as a covariate. Both analyses showed that ‘Gender’ did not significantly interact with our dependent measures (ER: all $F(1, 63) < .26$, all $p > .61$ and RT: all $F(1, 63) < .86$, all $p > .36$).

type of trial (congruent, incongruent). We performed generalised linear mixed models applying the Laplace Approximation method to estimate degrees of freedom and generate *p*-values for the AR analyses, and we run linear mixed models fitting with Maximum Likelihood Estimation and Satterthwaite's method for the RT analyses.

See section 2.2.2.3 for details on the software used to conduct all the different analyses.

2.4.3 Results

2.4.3.1 Mixed Analyses of Variance (ANOVA).

Outliers and omissions were excluded from the analyses (3.88% of the total trials). Four participants were removed from the analyses (1 colour-synaesthete, 1 sequence-synaesthete, and 2 non-synaesthetes) because they had at least one mean ER above 50% for any of the tasks' conditions. Overall, low ER were observed in all groups (non-synaesthetes: M = 3.26, SD = 3.03; colour-synaesthetes: M = 2.58, SD = 2.51; sequence-synaesthetes: M = 3.68, SD = 3.35). As expected, the ER analyses revealed a main effect for the factor 'Congruency' ($F(1, 64) = 73.13$, $p < .001$, $\eta_p^2 = .53$) with larger ER on incongruent (M = 5.59, SD = 5.2) than congruent trials (M = .76, SD = 1.05; $d = 1.05$) (Fig. 7A). A main effect of the factor 'Target-distractor' was also observed ($F(1, 64) = 54$, $p < .001$, $\eta_p^2 = .46$). In particular, ER were significantly larger for visual distractor (M = 5.67, SD = 5.58) than for visual target trials (M = .68, SD = .33; $d = .92$) (Fig. 7A).

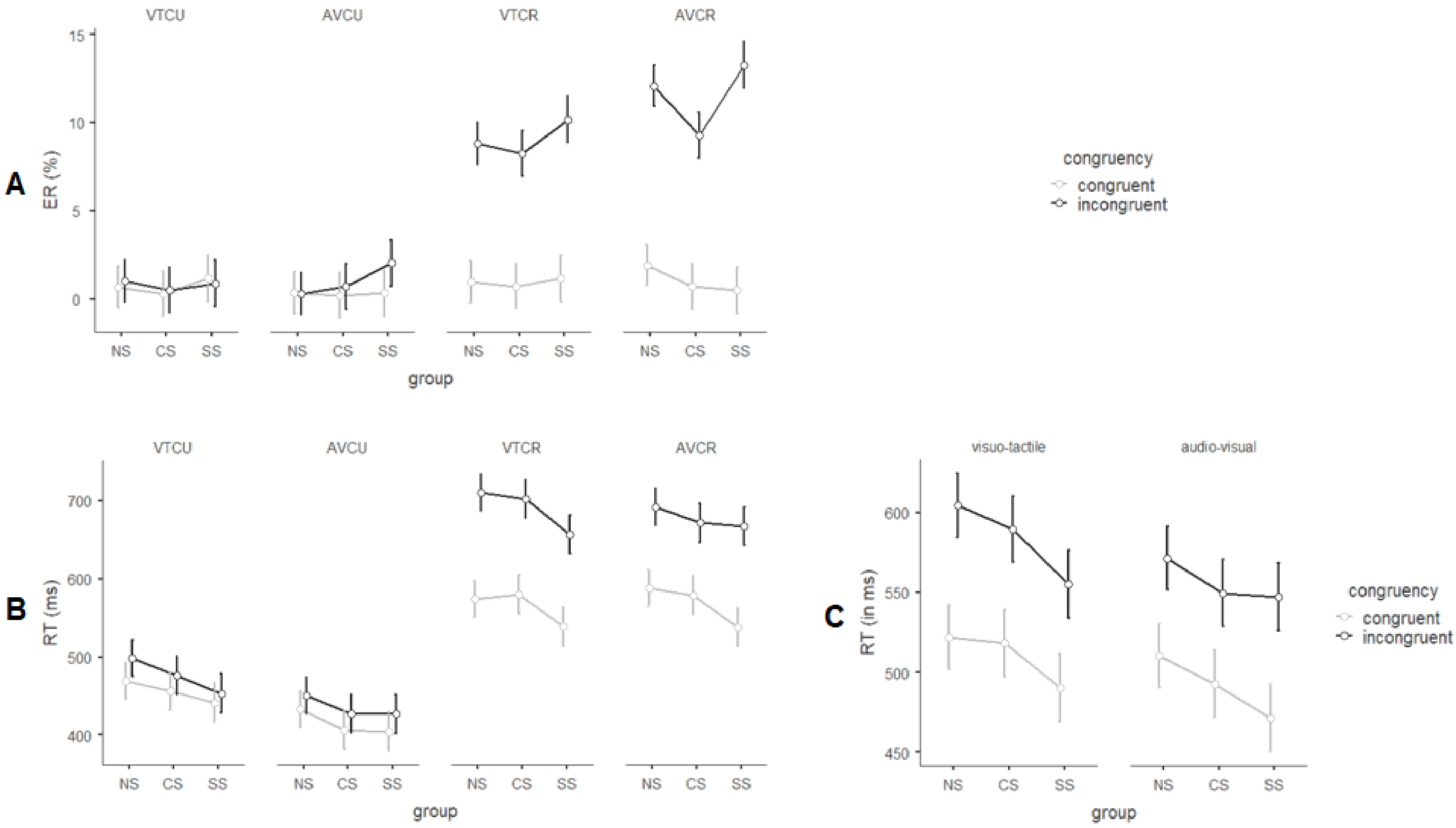


Figure 7. Mean error rates in (ER; Figure **A**) and mean reaction times in milliseconds (RT; Figure **B**) and their corresponding standard error means (error bars) for congruent and incongruent trials, for each task of Study 3 (Visuo-Tactile Concurrent-Unrelated – VTCU; Audio-Visual Concurrent-Unrelated – AVCU; Visuo-Tactile Concurrent-Related – VTCR; Audio-Visual Concurrent-Related – AVCR) and group (non-synaesthetes – NS; colour-synaesthetes – CS; sequence-synaesthetes – SS). Both the ER and RT analyses showed a main effect of 'Congruency' (both p < .001), reflecting expected congruency effects (CE; i.e. slower RT and higher ER for incongruent than congruent trials). In addition, the RT analysis revealed a three-way interaction between 'Congruency', 'Sensory modalities', and 'Group' (p = .034; Figure **C**). However, post-hoc tests comparing CE between groups separately for the combined visuo-tactile (VTCU and VTCR) and the combined audio-visual (AVCU and AVCR) tasks did not show though any statistical differences between groups.

We found interactions between 'Congruency' and 'Target-distractor' ($F(1, 64) = 60.3$, $p < .001$, $\eta_p^2 = .49$) and 'Congruency' and 'Sensory modalities' ($F(1, 64) = 5.74$, $p = .019$, $\eta_p^2 = .08$). Post-hoc comparisons determined that CE (i.e. incongruent minus congruent average ER) were significantly larger for visual distractor than visual target trials (M = .38, SD = 1.87 and M = 9.27, SD = 9.02, respectively; $Z(66) = 54.5$, $p < .001$, $d = .96$) (Fig. 7A). On the other hand, CE were significantly larger for audio-visual (M = 5.55, SD = 5.96) compared to visuo-tactile trials (M = 4.10, SD = 4.47; $Z(66) = 712$, $p = .012$, $d = .28$) (Fig. 7A). The factor of interest 'Group' and all the interactions involving it were not significant (all $F(2, 64) < 1.10$, all $p > .34$; Fig. 7A).

The RT analyses showed a main effect of ‘Congruency’ ($F(1, 64) = 374$, $p < .001$, $\eta_p^2 = .86$), with slower RT for incongruent (M = 570, SD = 94.3) than for congruent trials (M = 501, SD = 87.8; $d = 2.40$) (Fig. 7B). A significant main effect was also observed for the factor ‘Target-distractor’ ($F(1, 64) = 236$, $p < .001$, $\eta_p^2 = .79$, showing that RT were significantly slower when vision was the distractor modality (M = 625, SD = 119) compared to when vision was the target modality (M = 446, SD = 80.9; $d = 1.91$) (Fig. 7B). The factor ‘Sensory modalities’ was also significant ($F(1, 64) = 8.52$, $p = .005$, $\eta_p^2 = .12$). RT were slower in the visuo-tactile tasks (M = 547, SD = 97.1) than in the audio-visual tasks (M = 524, SD = 97.1; d =.36) (Fig. 7B).

In addition, the analyses showed interactions between the factors of ‘Congruency’ and ‘Target-distractor’ ($F(1, 64) = 193$, $p < .001$, $\eta_p^2 = .75$). Follow-up paired samples *t*-tests indicated that CE were significantly larger for visual distractor (M = 117, SD = 55.3) than for visual target trials (M = 20.5, SD = 13.2; $Z(66) = 1$, $p < .001$, $d = 1.72$) (Fig. 7B). There was also a ‘Target-distractor’ x ‘Sensory modalities’ interaction ($F(1, 64) = 6.33$, $p = .014$, $\eta_p^2 = .09$). RT were significantly slower for visuo-tactile trials compared to audio-visual ones when vision was the target modality (visuo-tactile: M = 466, SD = 94.2; audio-visual: M = 425, SD = 81.6; $t(66) = 4.84$, $p < .001$, $d = .59$), but not when vision was the distractor modality (visuo-tactile: M = 627, SD = 120; audio-visual: M = 623, SD = 137; $t(66) = .37$, $p = .72$) (Fig. 7B).

The main effect of interest ‘Group’ was not significant ($F(2, 64) = .91$, $p = .41$). However, the analysis revealed a three-way interaction between ‘Congruency’, ‘Sensory modalities’, and ‘Group’ ($F(2, 64) = 3.57$, $p = .034$, $\eta_p^2 = .10$) (Fig. 7C). To examine this interaction further, we run independent-samples *t*-tests comparing the CE between groups separately for visuo-tactile and audio-visual trials after previously checking the presence of significant CE for each task and group (see Table 5 for a summary of these control analyses). The analyses revealed no differences between groups for the CE of visuo-tactile trials combined (non-synaesthetes vs. colour-synaesthetes: $U(45) = 250$, $p = .63$; non-synaesthetes vs. sequence-synaesthetes: $U(44) = 221$, $p = .40$; colour- vs. sequence-synaesthetes: $t(39)$ =

.70, $p = .48$; Bonferroni-adjusted α for multiple comparisons = .017) (see the 'combined' values in Table 5 for the mean and standard deviation statistics of each group). There were no group differences either for the CE of audio-visual trials combined (non-synaesthetes vs. colour-synaesthetes: $U(45) = 265$, $p = .87$; non-synaesthetes vs. sequence-synaesthetes: $U(44) = 195$, $p = .15$; colour- vs. sequence-synaesthetes: $U(39) = 143$, $p = .08$; Bonferroni-adjusted α for multiple comparisons = .017). All other interactions involving the factor 'Group' were not significant (all $F(2, 64) < .91$, all $p > .41$).[19]

Table 5.
Congruency effects values and statistics for each task of Study 3, as well as the visuo-tactile and audio-visual tasks combined, by group.

Group	*N*	Task	CE	CE Statistics
Non-synaesthetes	26	VTCU	29 (25.6)	$F(1, 25) = 33.2$, $p < .001$, $\eta_p^2 = .57$
		VTCR	136 (85.2)	$F(1, 25) = 66.5$, $p < .001$, $\eta_p^2 = .73$
		VT combined	82.6 (44.6)	
		AVCU	17.6 (19.3)	$F(1, 25) = 21.5$, $p < .001$, $\eta_p^2 = .46$
		AVCR	104 (70.4)	$F(1, 25) = 57$, $p < .001$, $\eta_p^2 = .70$
		AV combined	60.9 (38.8)	
Colour-synaesthetes	21	VTCU	20 (21.3)	$F(1, 20) = 22.6$, $p < .001$, $\eta_p^2 = .53$
		VTCR	123 (57.4)	$F(1, 20) = 95.8$, $p < .001$, $\eta_p^2 = .83$
		VT combined	71.3 (30.4)	
		AVCU	21.3 (16.5)	$F(1, 20) = 34.9$, $p < .001$, $\eta_p^2 = .64$
		AVCR	92.2 (42.3)	$F(1, 20) = 99.8$, $p < .001$, $\eta_p^2 = .83$
		AV combined	56.7 (23.3)	
Sequence-synaesthetes	20	VTCU	11.5 (14.6)	$F(1, 19) = 12.4$, $p = .002$, $\eta_p^2 = .40$
		VTCR	118 (59.6)	$F(1, 19) = 78.7$, $p < .001$, $\eta_p^2 = .81$
		VT combined	64.8 (28)	
		AVCU	22.3 (13.9)	$F(1, 19) = 51.3$, $p < .001$, $\eta_p^2 = .73$
		AVCR	129 (87.2)	$F(1, 19) = 43.8$, $p < .001$, $\eta_p^2 = .70$
		AV combined	75.7 (41.2)	

Note: N = Sample size; Standard Deviations in parentheses; CE = Congruency effects (i.e. reaction time differences in milliseconds between incongruent and congruent trials); VTCU = Visuo-Tactile Concurrent-Unrelated; VTCR = Visuo-Tactile Concurrent-Related; AVCU = Audio-Visual Concurrent-Unrelated; AVCR = Audio-Visual Concurrent-Related; VT combined (VTCU and VTCR average); AV combined (AVCU and AVCR average).

[19] Since there were some issues with the assumption of normality of the dependent variables, we run additional non-parametric repeated-measures Friedman χ^2 tests. The analyses confirmed the main effects observed for ER regarding 'Congruency' ($\chi^2(1) = 59.2$, $p < .001$) and 'Target-distractor' ($\chi^2(1) = 48.5$, $p < .001$). The RT main effects were also ratified: Congruency' ($\chi^2(1) = 67$, $p < .001$), 'Target-distractor' ($\chi^2(1) = 67$, $p < .001$), and 'Sensory modalities' ($\chi^2(1) = 9.33$, $p = .002$).

2.4.3.2 Bayesian interference analyses.

The ER analyses provided strong evidence in favour of the null hypothesis regarding the interaction model (i.e. ‘Congruency’ x ‘Group) in the VTCU and AVCU tasks (BF_{10} = .036 and BF_{10} = .055, respectively). The ‘Congruency’-only models were the best performing ones with extreme evidence in the VTCR and AVCR tasks (VTCR: BF_{10} = 1.176e+9 and AVCR: BF_{10} = 7.680e+8). There was also extreme evidence in favour of the alternative hypothesis with respect to the interaction models (VTCR: BF_{10} = 2.086e+7; AVCR: BF_{10} = 2.909e+7). However, the exclusion of the interaction term was supported by substantial evidence (VTCR: Inclusion BF_{10} = .140; AVCR: Inclusion BF_{10} = .224). The RT analyses revealed that the ‘Congruency’-only models were the best performing ones for all tasks with extreme evidence (VTCU: BF_{10} = 1.282e+8; AVCU: BF_{10} = 2.023e+11; VTCR: BF_{10} = 1.797e+19; AVCR: BF_{10} = 9.550e+15). There was also extreme evidence in favour of the alternative hypothesis for the interaction models in all tasks (VTCU: BF_{10} = 1.626e+8; AVCU: BF_{10} = 1.741e+10; VTCR: BF_{10} = 1.616e+18; AVCR: BF_{10} = 1.926e+15). However, there was substantial evidence supporting the exclusion of the interaction term in the AVCU and VTCR tasks (Inclusion BF_{10} = .156 and Inclusion BF_{10} = .161, respectively) and the evidence was inconclusive for the VTCU and AVCR tasks (Inclusion BF_{10} = 1.893 and Inclusion BF_{10} = .415). The complete analyses can be consulted in Appendix B.

2.4.3.3 Linear mixed model analyses.

For the AR analyses, model fit values associated with fixed effects (marginal pseudo-R^2) were low or very low for all tasks (VTCU: marginal R^2 = .001; AVCU: marginal R^2 = .001; VTCR: marginal R^2 = .04; AVCR: marginal R^2 = .13). The model (i.e. AR ~ ‘Congruency’ x ‘Group’) was statistically significant for all tasks (all p < .001). ‘Congruency’ significantly predicted AR for the VTCR and AVCR tasks, with higher AR for congruent than incongruent trials (both p < .001); but not in the VTCU and AVCU tasks (p = .39 and p = .73, respectively). ‘Group’ did not predict AR or interacted with ‘Congruency’ for any of the tasks (all p < .21). The random effect of participant was a source of variation in all tasks' models and these were significantly preferred to analogous no-random effects models (all p < .001). However, conditional model fit values (conditional pseudo-R^2), which consider fixed and random effects, were only moderate for the AVCR task (VTCU: conditional R^2 = .009; AVCU: conditional R^2 = .009; VTCR: conditional R^2 = .08; AVCR: conditional R^2 = .27).

Regarding the RT analyses, marginal pseudo-R^2 were low for all tasks (VTCU: marginal R^2 = .02; AVCU: marginal R^2 = .01; VTCR: marginal R^2 = .10; AVCR: marginal R^2 = .07). The models were statically significant for all tasks (all p < .001). 'Congruency' significantly predicted RT for all tasks, with slower RT for incongruent compared to congruent trials (all p < .001). 'Group' did not predict RT for any of the tasks (all p > .29) and it did not interact with 'Congruency' for the AVCU, VTCR, and AVCR tasks (all p > .18). However, sequence-synaesthetes had a moderating effect on 'Congruency' on the VTCU task, these participants experiencing significantly smaller CE than non-synaesthetes (β = -21.1, SE = 5.98, t(71.4) = -3.53, p < .001; Fig. 8). A similar trend was observed between colour- and sequence-synaesthetes, but it did not reach significance (β = -12.1, SE = 6.31, t(71.1) = -1.92, p = .06). There were no differences between non-synaesthetes and colour-synaesthetes (β = 9, SE = 5.90, t(71.5) = 1.53, p = .13). The random effect of participant was an important source of variation in all tasks' models, which were significantly preferred to no-random effects models (all p < .001). Conditional pseudo- R^2 values were considerable for all tasks (VTCU: conditional R^2 = .46; AVCU: conditional R^2 = .48; VTCR: conditional R^2 = .47; AVCR: conditional R^2 = .55). [20] See Appendix C for detailed statistics.

[20] Visual inspection of residual plots revealed slight deviations from linearity, homoscedasticity, and normality. For that reason, we run the same models as generalised linear mixed models using the Penalised Quasi-Likelihood method (*MASS* package; Venables & Ripley, 2002). The analyses replicated the significance of 'Congruency' as a predictor of RT for all tasks, as well as the sequence-synaesthetes vs. non-synaesthetes and colour-synaesthetes differences in congruency in the VTCU; no other differences emerged. Generalised linear mixed models were already used to analyse AR, accounting thus for the non-normality of data.

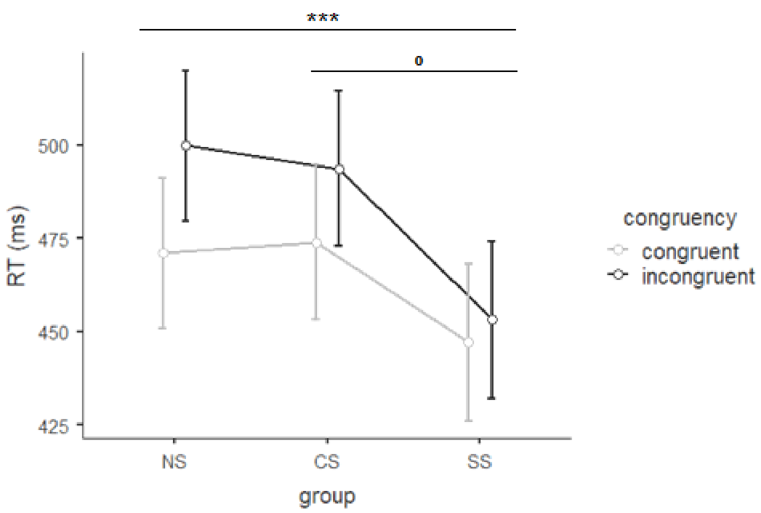


Figure 8. Mean reaction times (RT) in milliseconds and their corresponding standard error means (error bars) for the congruent and incongruent trials of the Visuo-Tactile Concurrent-Unrelated (VTCU) task of Study 3, for each group (non-synaesthetes – NS; colour-synaesthetes – CS; sequence-synaesthetes – SS). The analyses showed that sequence-synaesthetes had a moderating effect on 'Congruency'. Specifically, these participants experienced significantly smaller congruency effects (i.e. differences between incongruent and congruent trials) than non-synaesthetes ($p < .001$), and a similar trend was observed between sequence- and colour-synaesthetes **(**$p = .06$).

2.4.4 Discussion

Study 1 and Study 2 showed that synaesthetes who experience -visual synaesthetic types such as letters-colours or calendar-forms appear to be better than non-synaesthetes at filtering conflicting information coming from different sensory modalities. However, this ability seems to be limited to irrelevant stimuli that matches their synaesthetic concurrent modality (vision in this case). The present study had the aim to replicate these findings and address two follow-up important questions: (a) whether this advantage is consistent across different sensory modality target-distractor combinations, and (b) whether all types of -visual synaesthetes show the same abilities.

To investigate this, matched groups of colour-synaesthetes (i.e. subjects who experienced synaesthesias involving -colour as the concurrent), sequence-synaesthetes (i.e. subjects who experienced sequence-space synaesthesia forms) and non-synaesthetes performed several variations of the classic Cross-modal Congruency Task (CCT) that assessed two aspects: the processing of visuo-tactile vs. audio-visual stimuli and the processing of distractors that were related to synaesthetes' concurrents (visual) vs. distractors that were unrelated (tactile or auditory). Four tasks resulted from the combination of these two factors: visuo-tactile concurrent-unrelated (VTCU; vision-target, touch-distractor), audio-visual concurrent-unrelated (AVCU; vision-target, audition-distractor), visuo-tactile concurrent-related (VTCR; touch-target, vision-distractor), and audio-visual concurrent-related (AVCR; audition-target, vision-distractor). As in the CCT, all tasks presented simultaneous targets and distractors and participants were asked to focus on one sensory modality and ignore the other one. The extent to which distractors interfered with the processing of targets was regarded as a measure of participants intermodal attention abilities.

Strong congruency effects (CE) were observed for all tasks and analyses. That is, irrelevant distractors significantly slowed participants performance and caused them to make more errors on incongruent trials (target and distractor presented in opposite locations) compared to congruent trials (target and distractor in the same location). The analyses also revealed that error rates were significantly larger and reaction times slower for visual distractor than visual target trials. In addition, error rates were larger for audio-visual compared to visuo-tactile trials, and reaction times slower for visuo-tactile compared to audio-visual trials when vision was the target modality but not when vision was the distractor modality. Importantly, there was also a three-way interaction between ‘Congruency’ (congruent vs. incongruent), ‘Sensory modalities’ (visuo-tactile vs. audio-visual), and ‘Group’ (non-synaesthetes vs. colour-synaesthetes vs. sequence-synaesthetes). Sequence-synaesthetes appeared to be less affected than the other groups by the task-irrelevant stimuli in the visuo-tactile tasks; conversely, distractors impaired their performance more severely in the audio-visual tasks. However, post-hoc analyses comparing groups’ CE size in the visuo-tactile and the audio-visual tasks separately did not reveal any statistically significant differences.

Linear mixed model analyses revealed that there was an important influence of participants’ individual variability in all tasks. This appeared to be a critical factor, since the reaction time analyses showed that sequence-synaesthetes had a moderating effect on congruency on the VTCU task. In particular, this group experienced significantly smaller CE than non-synaesthetes and a similar trend was observed in relation to colour-synaesthetes, but it did not reach significance. Hence, these analyses seem to confirm and specify the above described differences observed for sequence-synaesthetes in the visuo-tactile tasks. The rest of the linear mixed model analyses confirmed the pattern of responses already observed. The Bayesian inference analyses partially supported these results. There was extreme evidence in favour of the alternative hypothesis for the interaction models (i.e. 'Congruency' x 'Group') in all tasks. However, there was substantial evidence supporting the exclusion of the interaction term in the AVCU and VTCR tasks and the evidence was inconclusive with respect to the VTCU and AVCR tasks. Moreover, the analyses determined that the models including only the congruency factor best explained the results in all tasks.

Based on previous studies, we hypothesised that synaesthetes would show advantages compared to non-synaesthetes at filtering distractors that matched their concurrents (i.e. vision), regardless of the sensory modality nature of the target stimuli with which these distractors were paired (i.e. advantages in the VTCR and AVCR tasks). Moreover, we predicted that no differences should be observed when the task-irrelevant stimuli were not visual and therefore did not match synaesthetes' concurrents (tactile or auditory distractors here; VTCU and AVCU tasks). However, here we observed that sequence-synaesthetes were better than non-synaesthetes *and* colour-synaesthetes when they had to ignore tactile distractors and focus on visual targets (VTCU task). Whilst this pattern of results appears to fail to replicate previous findings (the VTCR and VTCU tasks corresponded to Study's 1 CCT and Study's 2 rCCT tasks, respectively), the present study introduced a key new element: the separate groupings of sequence- and colour-synaesthetes. As observed here and already pointed in Study 1, sequence-synaesthetes seem to have a distinct attentional profile. However, in Studies 1 and 2, sequence- and colour-synaesthetes were not only mixed, but there were also quite a few participants who had both types (Study 1 N = 6, Study 2 N = 7). Thus, it might be possible that if colour-synaesthetes do not experience attentional advantages, having them in the sample could have had an influence on the results of those studies. Moreover, it is not clear whether people who present both types of synaesthesias have the same attentional profile shown by sequence- or colour-synaesthetes – or a different one altogether.

These both-synaesthete (N = 11[21]) participants were removed from the main analyses of the present study for these reasons, but we decided to run exploratory analyses comparing their performance to the rest of the groups. To do that, we conducted mixed analyses of variance (ANOVAs) of reaction times with the factors of 'Congruency' (congruent, incongruent) and 'Group' (non-, colour-, sequence-, both-synaesthetes), separately for each task. Visual examination of the means suggested that both-synaesthetes tended to show a similar profile to colour-synaesthetes and, thus, different to sequence-synaesthetes (Fig. 9). In the VTCU task, the results revealed a significant interaction ($F(3, 74) = 2.78$, $p = .047$, $\eta_p^2 = .10$), reflecting the identified differences between sequences-synaesthetes and non-synaesthetes. Follow-up independent samples t-tests indicated no differences between both-synaesthetes and the other groups (all $p > .23$, Bonferroni-adjusted $\alpha = .017$ – we omitted here all the comparisons already assessed in the main analyses). There were neither interactions or other significances of interest for the rest of the tasks (all $F(3, 74) < 1.18$, all $p > .32$; see Fig. 9 and Appendix E for the complete analyses).

[21] One both-synaesthete participant had to be removed from the exploratory analyses because they had at least one mean ER above 50% for any of the tasks' conditions.

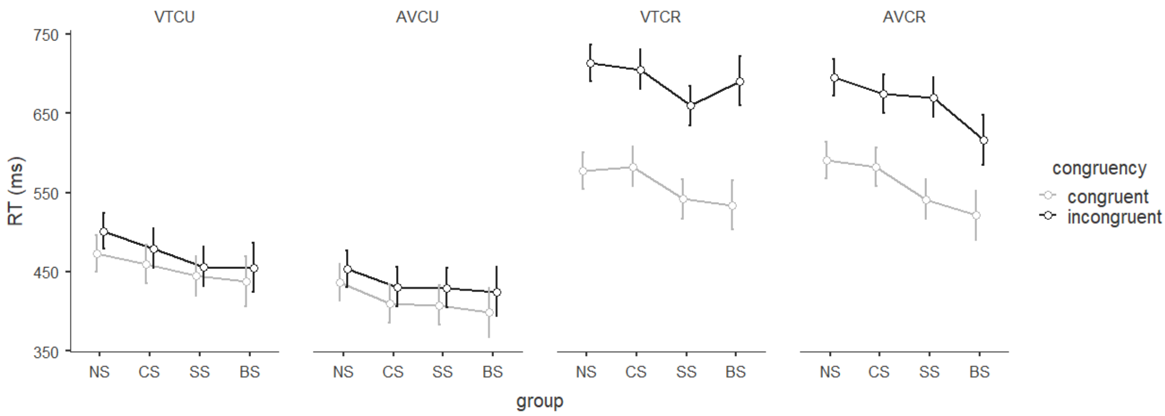


Figure 9. Mean reaction times in milliseconds (RT) and their corresponding standard error means (error bars) for congruent and incongruent trials, for each task of Study 3 (Visuo-Tactile Concurrent-Unrelated – VTCU; Audio-Visual Concurrent-Unrelated – AVCU; Visuo-Tactile Concurrent-Related – VTCR; Audio-Visual Concurrent-Related – AVCR) and group (non-synaesthetes – NS; colour-synaesthetes – CS; sequence-synaesthetes – SS; both-synaesthetes – BS). The analyses only revealed an interaction between 'Congruency' and 'Group' for the VTCU task (p = .047). However, post-hoc tests showed no differences between both-synaesthetes and the other groups (all p > .23).

Therefore, the presence or absence of both-synaesthetes does not seem to clarify the lack of group differences in Study's 3 VTCR task compared to Study's 1 CCT. Following the same rationale, if the synaesthetic advantage at ignoring visual distractors in the latter task is explained by the presence of different types of synaesthetes (colour-, sequence-, and both-synaesthetes) in the sample, we should observe here individual colour-, sequence-, and both-synaesthetes differences with respect to non-synaesthetes. All synaesthetes who participated in the different studies shared the experience of -visual concurrents, but most of them also experienced additional types of synaesthesias related or not to vision – and rarely a synaesthete had the same exact types of synaesthesias and/or experienced them in the same degree as another synaesthete. Thus, given the highly diverse nature of synaesthetes, it is likely that further sampling differences could be the cause of these results' incongruencies.

When screening for the participants of this study, we encountered a series of individuals who failed -colour and sequence-space consistency tests (at both threshold criteria) but reported having other types of synaesthesias (involving either -colour, -visual, and/or other types of concurrents) that could not be assessed [through consistency tests].

These subjects, who we called 'other-synaesthetes' (N = 22[22]), were removed from the main analyses because they did not fit our sample inclusion criteria, but we decided to run exploratory analyses evaluating their performance here in order to examine the scope of the differences between different types of synaesthetes. We conducted mixed analyses of variance (ANOVAs) of reaction times with the factors of 'Congruency' (congruent, incongruent) and 'Group' (non-synaesthetes, colour-synaesthetes, sequence-synaesthetes, other-synaesthetes), separately for each task. No group differences were observed for the AVCU, VTCR, or AVCR tasks (all $F(3, 85) < 1.53$, all $p > .21$; see Appendix E for the complete analyses). However, the analyses showed again a 'Congruency' x 'Group' interaction in the VTCU task ($F(3, 85) = 3.60$, $p = .017$, $\eta_p^2 = .11$). Post-hoc comparisons revealed that other-synaesthetes presented significantly smaller CE than non-synaesthetes (M = 12.8, SD = 19.6 and M = 29, SD = 25.6, respectively; $U(46) = 152$, $p = .005$, $d = .70$; Bonferroni-adjusted $\alpha = .017$ – we omitted here all the comparisons already assessed in the main analyses). Importantly, no differences were found between other-synaesthetes and colour- and sequence-synaesthetes (colour-synaesthetes: M = 20, SD = 19.2; $t(41) = 1.20$, $p = .24$ and sequence-synaesthetes: M = 11.5, SD = 14.6; $t(40) = .24$, $p = .81$) (Fig. 10). Therefore, the fact that other-synaesthetes showed no differences with both colour- and sequence-synaesthetes suggests that they might be have a different attentional profile with respect to other groups of synaesthetes.

[22] Two other-synaesthetes participants had to be removed from the exploratory analyses because they had at least one mean ER above 50% for any of the tasks' conditions.

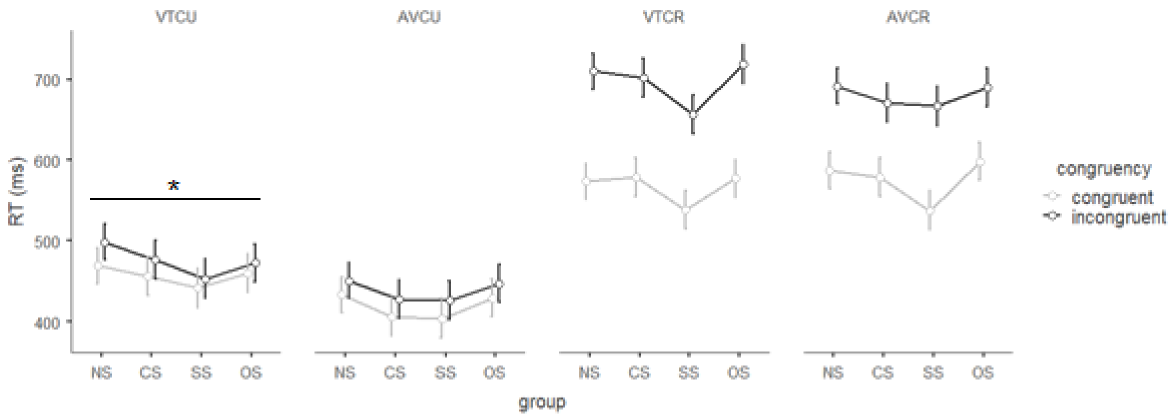


Figure 10. Mean reaction times in milliseconds (RT) and their corresponding standard error means (error bars) for congruent and incongruent trials, for each task of Study 3 (Visuo-Tactile Concurrent-Unrelated – VTCU; Audio-Visual Concurrent-Unrelated – AVCU; Visuo-Tactile Concurrent-Related – VTCR; Audio-Visual Concurrent-Related – AVCR) and group (non-synaesthetes – NS; colour-synaesthetes – CS; sequence-synaesthetes – SS; other-synaesthetes – OS). The analyses only revealed an interaction between 'Congruency' and 'Group' for the VTCU task (p = .017). Post-hoc tests showed that other-synaesthetes presented significantly smaller congruency effects (CE; i.e. differences in milliseconds between incongruent and congruent trials) than non-synaesthetes (p = .005); no differences in CE size were found between other- and colour- or sequence-synaesthetes (both p > .24).

In sum, the present study's results show a complex scenario. However, they seem to suggest the singularity of sequence-synaesthetes, not only with respect to non-synaesthetes but also to other types of synaesthetes – or, at least, other types of -visual synaesthetes such as those experiencing -colour percepts. This is of relevance because although there are many different types of synaesthetes known to date (Day, 2019), colour-synaesthetes are one of the most prevalent types and most studied groups, and are often taken as representative of all synaesthetes (e.g. Simner & Hubbard, 2013; Ward, 2013). The differences observed here suggest that that might not be the case and, therefore, a cautionary approach should be taken when assessing only colour-synaesthetes or when different types of synaesthetes are grouped together.

2.5 Chapter discussion

Studies 1 and 2 investigated whether -visual synaesthetes (i.e. those experiencing at least one synaesthesia type involving visual concurrents or triggers such as grapheme-colour or sequence-space synaesthesia) were better than-non synaesthetes at filtering out task-irrelevant stimuli in different conflict tasks. In Study 1, participants were asked to focus on tactile targets and ignore visual distractors (Cross-modal Congruency Task; CCT) in one task, and they were presented with visual targets and different types of visual distractors in another task (Flanker Task; FT) (see section 2.2.2.2). In Study 2, subjects performed a different version of the CCT in which the target-distractor sensory modalities were reversed and thus subjects had to attend to visual targets whilst supressing tactile distractors (reversed CCT; rCCT). In addition, participants completed a visual unimodal version of the CCT (visual Unimodal Conflict Task; vUCT) which presented different visual targets and distractors (see section 2.3.2.2). No group differences were observed between -visual synaesthetes and non-synaesthetes when they had to ignore tactile distractors (rCTT task) paired with visual targets or visual distractors in unimodal presentations in which the target was also visual (FT and vUCT tasks). In contrast, -visual synaesthetes (and particularly those who experienced sequence-space synaesthesias) were more efficient than non-synaesthetes at ignoring the visual distractors simultaneously presented with tactile targets (CCT task).

However, this finding was not replicated in a subsequent study. Study 3 assessed a new sample of participants with the CCT and the rCCT (here renamed Visuo-Tactile Concurrent-Related or VTCR task and Visuo-Tactile Concurrent-Unrelated or VTCU task, respectively). In addition, we evaluated whether the observed synaesthetic advantage was consistent across different sensory modalities combinations. To that aim, participants performed as well audio-visual versions of the same visuo-tactile tasks (Audio-Visual Concurrent-Related or AVCR task and Audio-Visual Concurrent-Unrelated or AVCU task). Lastly, we decided to investigate whether different types of -visual synaesthetes showed the same abilities by specifically comparing matched groups of non-synaesthetes, colour-synaesthetes (i.e. -visual synaesthetes experiencing only -colour associations such as months-colours), and sequence-synaesthetes (i.e. -visual synaesthetes experiencing only sequence-space synaesthesias such as number-forms).

Mixed Analyses of Variances showed a three-way interaction between the factors of ‘Congruency’ (congruent vs. incongruent), ‘Sensory modalities’ (visuo-tactile vs. audio-visual), and ‘Group’ (non- vs. colour- vs. sequence-synaesthetes), but post-hoc analyses comparing groups' CE size in the visuo-tactile and the audio-visual tasks separately did not reach significance. However, linear mixed model analyses, which take into account participants' individual variability, revealed that sequence-synaesthetes were better than non-synaesthetes *and* colour-synaesthetes at filtering tactile irrelevant distractors presented with visual targets (VTCU task); no other group differences were observed in the rest of the tasks. These analyses also determined that there was indeed an important influence of participants' individual variability in all tasks. This influence was also consistently observed throughout the different tasks assessed in Studies 1 and 2. Previous research has shown that individual differences in inhibition and conflict control are meaningful and should not only be considered a source of error variance (e.g. van den Wildenberg et al., 2010). Moreover, individual differences in these abilities have been observed to correlate with differences in both brain structure and function (e.g. Aron, Robbins, & Poldrack, 2014; Forstmann et al., 2008; Tabibnia et al., 2011). Lastly, it is worth mentioning that the three studies' results were largely supported by Bayesian interference analyses, providing substantial to extreme evidence in favour of the alternative hypothesis for models that fitted the results observed (i.e. high explicative power of the proposed 'Congruency' or 'Congruency' x 'Group models, depending on each task, in comparison to null models).

Given the specific differences observed for sequence-synaesthetes in Study 3, the analysis of different types of -visual synaesthetes together in Studies 1 and 2 could at least explain part of the discrepancies detected. In Study 3, we purposely did not include both-synaesthetes (i.e. those experiencing both colour- and sequence-space synaesthesias) in the main analysis in order to study differences between different types of synaesthetes, but we later run exploratory analyses comparing these individuals to the rest of the groups to see if these individuals could explain the discrepancies between the findings of Studies 1 and 3. In Study 3 we also omitted from the main analyses a group of people who failed both -colour and sequence-space consistency tests but reported having other types of synaesthesias (i.e. other-synaesthetes). We run pertinent exploratory analyses evaluating their performance with the aim to investigate the extent of differences between types of synaesthetes. No differences were found between both-synaesthetes and the other groups, but other-synaesthetes showed smaller CE than non-synaesthetes in the VTCU task and, importantly, they did not differ compared to colour- and sequence-synaesthetes.[23] Although all these exploratory analyses present small samples and thus might lack power, they reinforce the notion that investigating attention – and perhaps cognitive processes in general – in synaesthetes is a more intricate enterprise than initially devised. In particular, the different results observed indicate that synaesthetic individual differences (with respect to synaesthesia types) do not only seem to be of great relevance, but it might be necessary to consider them to understand other synaesthetic mental functions. Future studies are needed to further evaluate these factors independently and in relation to attentional processes.

Differences amongst synaesthetes might not only be qualitative (i.e. synaesthete or not). In our screening interviews, we observed that synaesthetes described experiencing synaesthesia in different degrees of intensity, frequency, stability, etc. (see sections 2.2.2.1.2, 2.3.2.1.2, and 2.4.2.1.2). These specifications could be considered different ways of defining synaesthesia strength and this variable might have a moderating effect on attention abilities. However, what is the exact definition of synaesthesia strength and how can it be measured? A first approach is to consider the number of synaesthesia types reported on the screening interview (Edinburgh Synaesthesia Screening Assessment or ESSA). Alternatively, the revised version of the ESSA used in Study 3 asked participants how much each type of synaesthesia applied to themselves, which allows the calculation of a mean overall degree of synaesthetic experience (see section 2.4.2.1.1 for further details). Lastly, we can also objectively measure synaesthesia strength through -colour and sequence-space consistency

[23] Although there were also other-synaesthetes in Study 1 and 2, we could not run exploratory analyses on them because, due to the screening process followed in those studies, they were intermixed with weak-synaesthetes (i.e. those passing synaesthetic consistency tests at the loose but not the strict threshold or those reporting to experience synaesthesias only 'Sometimes' (see section 2.2.2.1.1). In addition, these other-synaesthetes subsamples of Study 1 and 2, as the both-synaesthetes subgroups of these studies, were too small to conduct any meaningful analyses

scores (respectively obtained from the Synesthesia Battery/Multisense Consistency Test and the Sussex's Sequence-Spatial Synaesthesia Diagnostic Test).

To investigate this, we examined exploratory Pearson correlations (or Spearman rank correlations for ordinal variables; e.g. Salkind, 2010; Schober, Boer, & Schwarte, 2018) between the described synaesthesia strength measures and the congruency effects or CE (i.e. reaction time differences in milliseconds between incongruent and congruent trials) of those tasks for which group differences were observed (i.e. Study 1: CCT, Study 3: VTCU). Importantly, the analyses were only conducted on the synaesthetes subsamples since, by definition, non-synaesthetes do not experience any degree of synaesthetic experience and including them would have biased the results. In addition, we run separate analyses for the colour- and sequence-synaesthetes subsamples of Study 3 given the differences observed in the main analyses. In Study 1, we could only assess the relationship between CE size and number of synaesthesia types reported because participants completed the older version of the ESSA and, thus, we could not calculate the overall degree of synaesthetic experience; and we omitted the -colour consistency analysis because we considered that the colour-subsample was too small (N = 9) too produce any meaningful results. The analyses showed that CE size in the CCT and the number of synaesthesia types reported by synaesthetes was not related (r = -.010, p = .97; Bonferroni-adjusted α = .017). No relationships were found either between the number of synaesthesias, the overall degree of synaesthetic experience, or synaesthetic consistency scores and the CE size of the VTCU task of the colour- or sequence-synaesthetes subsamples (all r > .35, all p < .13; Bonferroni-adjusted α = .017) (see Appendix E for the complete analyses). Therefore, these results do not suggest that variations in synaesthetic strength between synaesthetes might influence synaesthetes' attention abilities. Or, at least, in the particular definitions of synaesthetic strength assessed here.

In order to try to understand the diverse results observed for the different studies conducted in this Chapter from another point of view, we combined the findings of these investigations pooling their effect sizes and reviewing them through a mini meta-analysis.

Given that only Study 3 compared different types of synaesthetes (i.e. colour- vs. sequence-synaesthetes), we omitted this comparison in the meta-analysis, focusing exclusively on non-synaesthete vs. synaesthete (colour-, sequence-, or both; depending on the study) differences. For the purposes of this meta-analysis, we calculated effect sizes with Hedges' *g* instead of Cohen's *d*, as Hedges' controls for the slight overestimating bias that has been observed in Cohen's *d* when used in small studies (Hedges, 1981). We run a random-effects-model (Borenstein, Hedges, Higgins, & Rothstein, 2011) to pool our effect sizes and perform the meta-analysis. In contrast to fixed-effects-model, random-effects-model assume that not only the effects of the individual studies deviate from the true effect of all studies due to sampling error, but that there is also another source of variance brought about the fact that the studies do not stem from a true single population (see Harrer, Cuijpers, Furukawa, & Ebert, 2019 for an in-depth explanation on the topic). There are several methods to estimate the variance of the distribution of true effect sizes (i.e. τ^2 or tau^2). The DerSimonian-Laird estimator (DerSimonian & Laird, 1986) is commonly used, but this method is prone to producing false positives, especially when the number of studies is small and there is substantial heterogeneity, common in the psychological field (e.g. Follmann & Proschan, 1999; Hartung, 1999; Hartung & Knapp, 2001; IntHout, Ioannidis, & Borm, 2014; Makambi, 2004). For that reason, we used the Hartunk-Knapp-Sidik-Jonkman (HKSJ) method, an alternative approach which is more conservative but produces more robust estimates (IntHout et al., 2014).

We performed meta-analyses with both the effect sizes derived from the results of the mixed ANOVAs and the linear mixed models. The analyses showed that the overall effect size estimates using the HKSJ method were low: g = -.12 (confidence interval g = -.38 to .15) for the mixed ANOVAs and g = -.085 (confidence interval g = -.35 to .18) for the linear mixed models. The prediction intervals of the models were also large: mixed ANOVAs (g = -.90 to 0.67) and linear mixed models (g = -.87 to 0.70) (see Fig. 11A and 11B). We also evaluated the between-study heterogeneity through three different metrics: the τ^2 (the direct variance in our meta-analysis; see above), the Cochran's Q^2 (the difference between the observed effect sizes and the random-effect-estimate of the effect size; Harrer et al., 2019); and the Higgins & Thompson's I^2 (percentage of variability in the effect sizes which is not caused by sampling error, derived from Q^2; Higgins & Thompson, 2002). In accordance with the low overall effect sizes estimates observed, the results revealed between-study heterogeneity. For the mixed ANOVA, we observed a $\tau^2 = .11$ (confidence interval $\tau^2 = 0$ to .41), a $I^2 = 42.4\%$ (confidence interval $I^2 = 0$ to 70.7%), and a $Q^2 = 19.1$ (close to significant: $p = .06$). The linear mixed models' analysis showed a very similar pattern: $\tau^2 = .33$ (confidence interval $\tau^2 = 0$ to .64), $I^2 = 43.9\%$ (confidence interval $I^2 = 0$ to 71.4%), and $Q^2 = 19.6$ ($p = .06$).

In addition, we performed Influence Analyses, which allow to detect whether there are any studies which influence the overall estimate of our meta-analyses through leave-one-out methodology. In both the mixed ANOVA and the linear mixed models, we can observe that Study's 1 CCT task (synaesthetes vs. non-synaesthetes) and Study's 3 VTCU task (sequence- vs. non-synaesthetes), which were the two studies in which we detected group differences, were the most influential (see Fig. 12A and 12B). None of the different metrics display any study in red, which marks influential cases according to Viechtbauer and Cheung's (2010) proposed interpretation guide. However, the authors themselves acknowledge the arbitrariness of this visual threshold, emphasising the need to interpret these results considering the overall context. In the present analysis, we consider that the different metrics ratify the results observed thus far: the CCT and the VTCU tasks' effect sizes are clearly influential, inducing the detected between-study heterogeneity as well. However, their influence might not be as high as to justify their exclusion from the pooled effect sizes considered. In sum, the mini meta-analyses conducted cannot explain the differences observed across different studies.

(All analyses were performed in R 3.5.1 – R Core Team, 2018 – with the following packages: *meta* – Balduzzi, Rücker, & Schwarzer, 2019; *metafor* – Viechtbauer, 2010; and *dmetar* – Harrer et al., 2019).

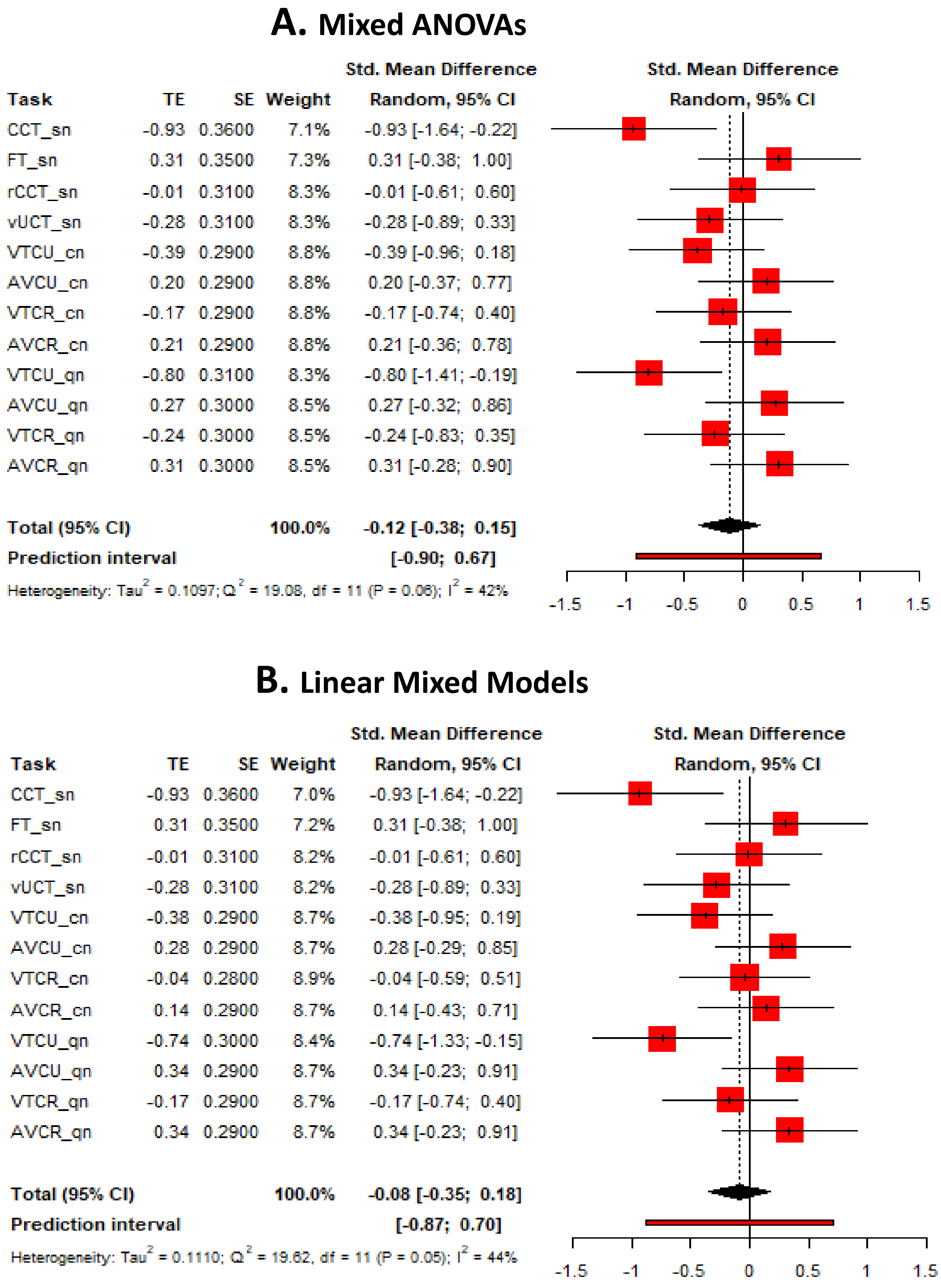


Figure 11. Output of the random-effects-model meta-analysis (Hartnuk-Knapp-Sidik-Jonkman method and Hedges' *g* estimator) of the different studies of Chapter II, for the mixed ANOVAs (**A**) and the linear mixed models results (**B**). CCT = Cross-modal Congruency Task (Study 1), FT = Flanker Task (Study 1), rCCT = reversed Cross-modal Congruency Task (Study 2), vUCT = visual Unimodal Congruency Task (Study 2), VTCU = Visuo-Tactile Concurrent-Unrelated (Study 3), AVCU = Audio-Visual Concurrent-Related (AVCU), VTCR = Visuo-Tactile Concurrent-Related, AVCR = Audio-Visual Concurrent-Related, sn = synaesthetes vs. non-synaesthetes, cn = colour- vs. non-synaesthetes, qn = sequence- vs. non-synaesthetes, TE = Effect Size, SE = Standard Error of the Effect Size, (Q^2 = Cochran's Q, I^2 = Higgins and Thompson's I^2.

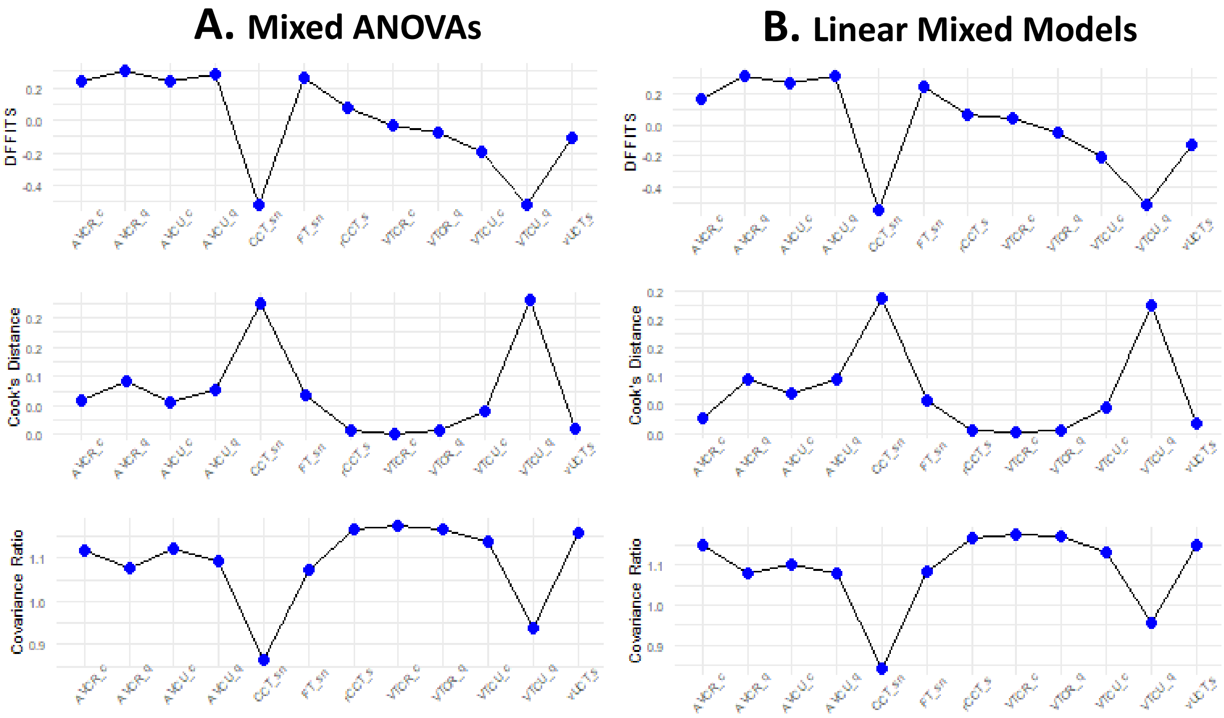


Figure 12. Output of the Influence analyses performed for the random-effects-model meta-analyses (Hartnuk-Knapp-Sidik-Jonkman method and Hedges' *g* estimator) of the different studies of Chapter II, for the mixed ANOVAs (**A**) and the linear mixed models results (**B**). DIFFITS = calculates how much the predicted pooled effect changes after excluding a study, in standard deviations; Cook's Distance = distance between the value once the study is included compared to when it is excluded; Covariance Ratio = determinant of the variance-covariance matrix of the parameter estimates when the full dataset is considered – values < 1 indicate that removing the study will lead to a more precise effect size estimation (i.e. less heterogeneity).

These studies represent a first attempt to investigate synaesthetes' intermodal attentional abilities. One relevant question which remains to be explored concerns the exact mechanisms responsible for the differences observed between synaesthetes and controls. Does synaesthetes' constant need to ignore their automatic synaesthetic percepts cause a 'synaesthetic attentional training'? Or is synaesthetes' attentional profile intrinsically distinct from the general population? Changes over the life span in synaesthetes' intermodal attentional abilities might support the 'training' hypothesis. In particular, one might expect that these attentional abilities depend on the amount of synaesthetic interference to which synaesthetes are exposed. Older synaesthetes, who have experienced irrelevant percepts for a longer period of time, should be better than younger ones at filtering out irrelevant information. According to this hypothesis then, a negative correlation should be expected between age and CE. While the age range of synaesthetes in the our studies were too narrow (18 to 31 years old) to provide meaningful insight into this question, it is worth noting that there is evidence showing that the number of audio-visual double-flash illusions experienced by synaesthetes is negatively correlated with age (Neufeld et al., 2012). Furthermore, evidence from the other end of the age spectrum shows that children with grapheme-colour synaesthesia experience difficulties in numerical tasks due to cognitive interference caused by digits presented in colours incongruent to their synaesthetic associations (Green & Goswami, 2008). These findings might suggest that synaesthetes attentional abilities are improved over time in a use-dependent fashion. That is, synaesthetes learn to deal with their synaesthetic concurrents. However, indirect evidence from associative learning studies in synaesthetes casts some doubts about the 'training' hypothesis (Bankieris and Aslin, 2016a; 2016b). In these studies, while synaesthetes performed better than non-synaesthetes in an explicit associative learning task, they seemed to experience greater interference during an implicit associative learning task. If synaesthetes learn to ignore their synaesthetic concurrents, they should be particularly able to train their attentional systems to ignore other irrelevant information (i.e. they should be less affected by interference). Future studies in this area are necessary to further evaluate these alternative hypotheses.

Finally, it is relevant to note that all synaesthetes tested in these studies primarily experienced -visual synaesthesias and, particularly, grapheme-colour and sequence-space synaesthesias. Considering the differences observed for different types of synaesthetes, it would be particularly interesting to further assess different types of synaesthetes with other types of -visual synaesthesias and, especially, synaesthetes with concurrents beyond the visual domain, much less studied in general in the synaesthetic research field. In addition, all tested synaesthetes were associators. Whilst associators experience their synaesthetic concurrents 'in the mind's eye', projectors report these experiences in external space (Dixon et al., 2004). There is contradictory evidence regarding behavioural advantages of projector over associator synaesthetes (e.g. Dixon et al., 2004; Rothen & Meier, 2009; Ward, Li, Salih,

& Sagiv, 2007) or about the existence of neural differentiation between the two groups (e.g. Rouw & Scholte, 2007; 2010). Nonetheless, different synaesthetic experiences could potentially imply different attentional processing strategies: Do projectors' ostensibly stronger synaesthetic interference cause stronger filtering abilities compared to associators? Or does this special synaesthetic experience make them less successful at ignoring their percepts and, hence, they have weaker filtering abilities? Future studies assessing different types of synaesthetes and different types of stimuli should clarify all these points.

3. Chapter III: Personality Trait Differences Between Synaesthetes (Study 4)

3.1 Introduction

Previous evidence has shown that individuals who experience synaesthesia seem to have a distinct personality profile compared to non-synaesthetes. Synaesthesia has been consistently linked to higher rates of Openness to Experience (Banissy et al., 2013a; Chun & Hupé, 2016; Rouw & Scholte, 2016 – but see Ward et al., 2018a); of Absorption/Fantasising, a dimension of empathy (Banissy et al., 2013a; Chun & Hupé, 2016; Rader & Tellegen, 1987; Rouw & Scholte, 2016); and of positive schizotypy (Banissy et al., 2012; Janik McErlean & Banissy, 2016). However, evidence is inconclusive for other personality traits. Banissy et al. (2013a) and Rouw and Scholte (2016) found that synaesthetes experienced higher rates of Neuroticism, but other studies did not observe differences in the scores of this trait between synaesthetes and controls (Chun & Hupé, 2016; Ward et al., 2018a). Similarly, Banissy and Ward (2007) and Rouw and Scholte (2016) observed that synaesthesia was associated with higher rates of Emotionality, but this difference has not been replicated in other investigations (Banissy et al., 2013a; Baron-Cohen, Robson, Lai, & Allison, 2016). Lastly, lower rates of Agreeableness (Banissy et al., 2013a) and Conscientiousness (Rouw & Scholte, 2016) have also been reported, but these results were not observed in other studies (Agreeableness: Chun & Hupé, 2016; Rouw & Scholte, 2016; Ward et al., 2018a; Conscientiousness: Banissy et al., 2013a; Chun & Hupé, 2016; Ward et al., 2018a).

One possible reason that might explain at least in part these inconsistencies is related to sampling differences. Most of the research on personality and synaesthesia has focused on a single type of synaesthesia, namely grapheme-colour synaesthesia. However, researchers in the field are becoming increasingly aware about the complex reality of synaesthesia and, in particular, about the fact that experiencing only one type of synaesthesia might be more the exception than the norm (e.g. Cytowic & Eagleman, 2009; Niccolai et al., 2012; Novich et al., 2011; Sagiv et al., 2006b). Therefore, although the majority of synaesthetes appear to have different types of synaesthesia, most synaesthesia studies (including the individual differences ones considered here) focus on one specific type and often neglect to report the presence (or absence) of additional types.

This is especially important if we consider that there is some initial evidence indicating that all synaesthetes might not present the same personality profile. Rouw and Scholte (2016) compared the Big Five personality traits and the emotional style of people experiencing coloured sequences, coloured sounds, coloured sensations, spatial sequences, non-visual synaesthesias, and sequence personalities. The authors observed that synaesthetes had higher scores than non-synaesthetes in the subscales of Openness to Experience, Neuroticism, Fantasising, and Emotionality, and lower scores for Conscientiousness. However, they did not find differences for these traits between the different synaesthetic groups. On the other hand, a recent study which focused exclusively on sequence-space synaesthetes showed that this group did not differ from controls in their scores for any of the Big Five personality traits, including Openness to Experience, Neuroticism, and Conscientiousness (Ward et al., 2018a). Thus, it is not clear whether sequence-synaesthetes experience the same personality profile as other synaesthetes (or specifically as grapheme-colour synaesthetes, since they have been the most studied group).

One critical difference between these two studies is that Ward and colleagues (2018a) assessed sequence-synaesthetes through a recently developed synaesthetic consistency test for this type of synaesthesia. Consistency tests objectively quantify how consistent synaesthetic associations are for a given person. For instance, in the sequence consistency test, subjects are asked to place the different months of the year, days of the week, and digits 0 to 9 in a blank computer screen over repeated trials. The less the distance between the different trial locations given for a particular month, day of the week, or digit (or, in other words, the more consistent a person is at giving the same location for each item), the stronger the synaesthetic association is regarded (see section 2.2.2.1.1 for further details). Although these tests are considered the 'gold-standard' of synaesthesia assessment, only a few of the personality studies presented here have used much more established -colour consistency tests (based on the same premises) and not all which have, have done it in a systematic way (e.g. only part of the sample).

This poses a problem that could account for some of the conflicting results found to date. First, it implies that is difficult to ascertain whether previous studies have assessed true synaesthetes and which types of synaesthesia did the participants exactly experienced. This might have resulted in different studies comparing different types of synaesthetes. Second, most studies tend to recruit participants via self-referral, a method which involves sampling biases. Simner et al. (2006) specifically identified differences in terms of synaesthetic prevalence and gender ratios between samples of self-referred synaesthetes and samples with participants recruited randomly (i.e. with no specific mention of synaesthesia). Carmichael et al. (2015) also suggested that self-referrers might have particular characteristics and, therefore, not be entirely representative of the synaesthetic population at large. Moreover, several studies have observed that self-referral and volunteer participants in experimental studies in general score higher in the personality traits of Extraversion, Agreeableness, and Openness to Experience (e.g. Brüggen & Dholakia, 2010; Dollinger & Leong, 1993; Lönnqvist et al., 2007; Marcus & Schütz, 2005). Thus, it is possible that some of the higher rates in personality traits observed for synaesthetes in previous studies respond in part to this bias.

Taken together, this initial evidence seems to question the assumption that all synaesthetes share a general personality profile. The aim of the present study is to directly investigate whether different types of synaesthetes might present specific personality profiles. In order to address this question, we will compare non-synaesthete controls, colour-synaesthetes (i.e. subjects who experienced synaesthesias involving -colour as the concurrent, e.g. grapheme-colour synaesthesia), and sequence-synaesthetes (i.e. subjects who experienced sequence-space synaesthesia forms; e.g. calendar-forms). We focused in these two groups of synaesthetes for three main reasons: (a) they are prevalent forms of synaesthesias amongst synaesthetes, (b) (because of that) they have been largely investigated in previous studies and we can therefore draw comparisons with previous findings, and (c) both types can be assessed through objective synaesthetic consistency measures.

All participants were assessed and compared on different personality questionnaires measuring the Big Five personality traits (Big Five Inventory), empathy (Interpersonal Reactivity Index), and schizotypy (Oxford-Liverpool Inventory of Feelings and Emotions). Given the literature reviewed, we predict that both colour- and sequence-synaesthetes will show higher scores than controls for the following traits: Openness to Experience (measured with Big Five Inventory), Fantasising (Interpersonal Reactivity Index), and positive schizotypy traits (Oxford-Liverpool Inventory of Feelings and Emotions). In addition, we further hypothesise that colour-synaesthetes will show higher rates for these personality traits than sequence-synaesthetes.

A secondary aim of the present study was to examine whether personality trait scores were modulated by differences in synaesthetic strength. The concept of synaesthetic strength is a matter of ongoing theoretical debate, but it has been operationalised in different ways. For example, by the number of synaesthesia types experienced (Havilk et al., 2015; Rouw & Scholte, 2016; Spiller et al., 2015). Rouw and Scholte (2016) correlated the number of synaesthesia types and personality scores and found that, for some of the traits that differentially characterise synaesthetes from controls such as Openness to Experience or Fantasising, there were significant positive relationships (i.e. the more the synaesthesia types experienced, the higher the scores on those personality traits). Synaesthetic strength can also be operationalised in terms of synaesthetic consistency scores. Hossain et al. (2018) observed that the amount of saturation of synaesthetic colours reported by grapheme-colour synaesthetes was positively correlated to the rates of Openness to Experience of these participants. However, scores on personality traits were not directly related to -colour consistency scores. In addition, synaesthetic strength can also be measured as the overall degree of (self-reported) synaesthetic experience in screening interviews or questionnaires such as the Edinburgh Synaesthesia Screening Assessment (ESSA), developed in this thesis (see Chapter IV for details).

We examined how these different operationalisations of synaesthesia strength correlated with those personality scores showing group differences. Given previous evidence, we expect to find positive relationships between personality rates and number of synaesthesia types, but not synaesthetic consistency scores. We have no predictions with respect to the overall degree of synaesthetic experience (as assessed by the ESSA screening assessment).

3.2 Methods

3.2.1 Personality questionnaires

The following described personality questionnaires were chosen due to their proven reliability and extensive usage, their relative brevity, and the fact that they have been previously administered in several of the previous studies investigating personality and synaesthesia.

- Big Five Inventory (BFI; John & Srivastava, 1999): Is a well-established measure of the Big Five personality traits, one of the most widely used models of personality and which has shown good reliability across age, culture, and time (at an individual level) (e.g. McCrae and Costa, 1987). The BFI has high validity against the NEO-Five Factor Inventory (NEO-FFI; Costa & McCrae, 1989) and Goldberg's (1993) mini-markers. It has 44 items distributed in five subscales: Extraversion (8 items; i.e. refers to the sociability and energy of the individual towards the outer world), Agreeableness (9 items; i.e. tendency for altruism and compliance); Neuroticism (8 items; individual's level of psychological distress, vulnerability, and self-consciousness); Conscientiousness (9 items; i.e. degree of motivation, competence, persistence, and self-discipline in goal-directed behaviour); and Openness to Experience (10 items; reflects how curious, excitable, and imaginative an individual is). Questions are rated on a 5-point Likert scale ranging from 1-'Disagree strongly' to 5-'Agree strongly' and averaged, higher scores reflecting higher expressions of the traits.

- Interpersonal Reactivity Index (IRI; Davis, 1980): The IRI is a questionnaire widely used to measure trait empathy. It comprises four subscales: Perspective Taking (i.e. capacity to adopt someone else's point of view), Fantasising (i.e. tendency to get involved in a film, novel, etc.), Empathic Concern (i.e. ability to show concern or feel sorry for other people's pain), and Personal Distress (i.e. feelings of anxiety caused by others' worry). The assessment contains 28 items (7 each subscale) measured on a 5-point Likert scale ranging from 0-'Does not describe me well' to 4-'Describes me very well'. Total scores per subscale can go from 0 to 28, higher scores reflecting heightened empathy for all the scales except for Personal Distress, which indicates self-oriented emotional reactivity.

- Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE; Mason & Claridge, 2006): It measures schizotypy or sub-clinical psychosis-alike symptoms exhibited in varying degrees by the general population. The questionnaire has 301 questions divided into the following independent subscales: Introvertive Anhedonia (27 items), Impulsive Non-conformity (23 items), Unusual Experiences (30 items), and Cognitive Disorganisation (24 items). Due to the lengthiness of the full questionnaire and given that differences between synaesthetes and non-synaesthetes have only been found for the Unusual Experiences and Cognitive Disorganisation traits (Banissy et al., 2012; Janik McErlean & Banissy, 2016), we administered only these subscales. Unusual Experiences measures positive symptomatology and refers to the proneness to apply aberrational and magical interpretations to the world. On the other hand, Cognitive Disorganisation taps into dysfunctional thought disorder and inattentive thinking. The questionnaire has a dichotomous forced-choice response where 'Yes' is scored as 1 and 'No' as 0, higher schizotypal traits corresponding to total average scores closer to 1.

3.2.2 Participants (Samples A and B)

The participants of the present study were pooled from the data collection process of Study 3 (Sample A) and a new data collection process shared across Studies 4 and 5 (Sample B). The project for Study 3/Sample A involved three different sessions in which 359 total individuals participated in all or part of these sessions. Subjects included here are those who completed the first and the second sessions. The first session consisted of answering on-line the revised version the Edinburgh Synaesthesia Screening Assessment (ESSA; see Chapter IV) and several personality questionnaires. The second session took place in the Lab and participants were interviewed with the same screening assessment[24] and completed a series of synaesthesia questionnaires and tests (including consistency tests for letters-colours and sequence-spatial synaesthesias). A total of 156 people met these criteria: 31 were classified as non-synaesthetes (i.e. participants failing both consistency tests and not reporting any synaesthetic experiences) and 86 as synaesthetes (i.e. participants passing either or both consistency tests; in particular there were 41 colour-synaesthetes, 18 sequence-synaesthetes, and 27 synaesthetes who presented both types). In addition, 39 more subjects were classified as other-synaesthetes (i.e. participants failing both consistency tests but reporting to have other types of synaesthesias that could not be assessed). See section 2.4.2.1 for the complete details.

Both in the present study (Study 4) and in Study 5 we used the loose synaesthetic consistency criteria (see section 2.2.2.1.1) to classify participants with the aim to capture as much as possible the complex reality and variability of the synaesthetic population, as we considered these key factors on the two topics covered (group and individual differences in personality traits and development of a screening tool). People experiencing both types of synaesthesias were, however, not included in the current study in order to properly examine differences between types of synaesthetes. In addition, other-synaesthetes were also removed because they did not fit our sample inclusion criteria. Thus, Sample A was finally composed of 31 non-synaesthetes, 41 colour-synaesthetes, and 18 sequence-synaesthetes. Table 6 shows the main descriptive statistics for the Study 4 final samples and Table 7 a

[24] Sixteen synaesthete participants (12 colour- and 4 sequence-synaesthetes) were not interviewed with the ESSA in the second session due to timing constraints related to other experimental projects that were being conducted in parallel. However, they were included in the present study's sample because all these participants met the independent requirement of passing the -colour or sequence consistency tests and, therefore, their synaesthetic status was well-established.

summary of the average consistency scores obtained for all the different groups and samples [25].

Table 6.
Descriptive, chi-square (χ^2), and t-*statistics of Study 4 groups' demographics.*

Demographics	Colour-synaesthetes	Sequence-synaesthetes	Non-synaesthetes	Statistics
Sample A				
N (male)	34 (7)	15 (3)	20 (11)	$\chi^2(2) = 3.90, p = .14$
Age (SD)	21.4 (2.74)	21.8 (1.62)	22.5 (3.06)	$F(2, 87) = 1.58, p = .21$
Handedness (left, ambidextrous)	35 (4, 2)	15 (1, 2)	29 (2, 0)	$\chi^2(4) = 3.77, p = .44$
Nº of (native) languages* (SD)	1.27 (.50)	1.39 (.61)	1.13 (.34)	$F(2, 87) = 1.78, p = .17$
Level of education** (SD)	2.49 (.60)	2.50 (.62)	2.42 (.56)	$F(2, 87) = .16, p = .86$
Sample B				
N (male)	41 (11)	9 (1)	111 (23)	$\chi^2(2) = .85, p = .65$
Age (SD)	21.1 (4.13)	20.6 (2.70)	21 (5.25)	$F(2, 193) = .0478, p = .95$
Handedness (left, ambidextrous)	47 (3, 2)	7 (3, 0)	117 (17, 0)	$\chi^2(4) = 10.5, p = .033$
Nº of (native) languages* (SD)	1.02 (.14)	1.20 (.42)	1.12 (.35)	$F(2, 193) = 2.52, p = .083$
Level of education** (SD)	2.29 (.57)	2.60 (.84)	2.17 (.51)	$F(2, 193) = 3.24, p = .041$

Note: *N* = Sample size, SD = Standard Deviation.
* Nº of (native) languages: 1 = Monolingual, 2 = Bilingual, 3 = Polylingual.
** Level of education: 1 = High School, 2 = Undergraduate, 3 = Master, 4 = PhD, 5 = Postdoc.

Table 7.
-Colour and sequence synaesthetic consistency scores of Study 4 participants, by sample and group.

Synaesthetic consistency measures	Colour-synaesthetes	Sequence-synaesthetes	Both-synaesthetes	Other-synaesthetes	Non-synaesthetes
Sample A					
-Colour SB/MCT consistency score (SD)*	.93 (.32)	1.99 (.40)	.92 (.32)	2.18 (.40)	2.34 (.46)
Sequence SDT consistency score (SD)	.69 (.86)	.12 (.08)	.10 (.07)	.70 (1)	.60 (.86)
Sequence SDT questionnaire score (SD)[1]	30.1 (6.94)	18.4 (3.99)	18.7 (3.64)	30.8 (6.89)	37.5 (5.54)
Sample B					
-Colour MCT consistency score (SD)	89 (7.10)	59.1 (11.6)	-	53.9 (10.6)	51.7 (10.3)
Sequence SDT consistency score (SD)	.39 (.91)	.05 (.02)	-	.84 (1.33)	.82 (1.36)
Sequence SDT questionnaire score (SD)[2]	33.6 (7.67)	14.9 (3.63)	-	28.4 (7.91)	34.8 (7.18)

Note: SD = Standard Deviation, SB = Synesthesia Battery (synaesthetic threshold < 1.43), MCT = Multisense Consistency Test (synaesthetic threshold ≥ 75%), SDT = Sussex's Sequence-Spatial Synaesthesia Diagnostic Test = SDT (consistency threshold < .300; questionnaire [1]< 25, [2]< 19 – see Footnote 14).
* Scores of participants of Sample A who completed the MCT were transformed to SB scores for homogenisation purposes.

[25] Power analyses were performed for sample size estimation based on the results of previous similar studies that assessed personality traits in synaesthesia. We selected those investigations that used the same personality measures as the present study and that compared synaesthetes and non-synaesthetes. In particular, Banissy et al. (2013a) was used to calculate estimations regarding BFI Openness to Experience and IRI Fantasising, and Banissy et al. (2012) and Janik McErlean and Banissy (2016) for O-LIFE Unusual Experiences. With error probability α = .05 and target power = .80, the projected total sample was *N* = 40 (±17 synaesthetes and ±23 controls) for BFI Openness to Experience, *N* = 84 (±35 synaesthetes and ±40 controls) for IRI Fantasising, and *N* = ±44 (±22 per group) for O-LIFE Unusual Experiences. Thus, our proposed samples should be adequate for the aims of this study: 41 colour-synaesthetes, 18 sequence-synaesthetes, and 31 non-synaesthetes (Sample A); 52 colour-synaesthetes, 10 sequence-synaesthetes, and 134 non-synaesthetes (Sample B).

The data collection of Sample B followed a similar procedure, but it was conducted exclusively on-line. Participants completed the revised ESSA, the personality questionnaires, and the synaesthetic consistency tests for letters-colour and sequence-space experiences[26]; the study lasting between 60 to 90 minutes in total. In order to be able to control any possible response biases due to the order of the assessments administered, half of the participants answered the ESSA first and then the personality questionnaires and the other half vice-versa[27]. However, all subjects completed these assessments before the synaesthetic consistency tests, so they were as naïve as possible to the ESSA and the aims of the study. As in Study 3/Sample A recruitment, the participants' call invited people to participate in a "perception experiment" and the word synaesthesia or its definition was purposely not present in the ad description or throughout the study instructions. Trying to match as much as possible with the demographic characteristics of Study 3/Sample A participants, all Sample B subjects were between 18 and 40 years-old, currently living in the UK, proficient at English, and had normal or corrected-to-normal vision. In addition, we stablished the criteria of having to use a laptop or desktop computer to complete the study (important to optimally do the consistency tests) and not having participated in any of the researcher's previous studies. Twenty-six participants had to be dropped due to not meeting (totally or partially) these criteria.

Two-hundred and seventy-five people met all the criteria and fully completed the described assessment measures: 134 non-synaesthetes, 62 synaesthetes (52 colour-synaesthetes and 10 sequence-synaesthetes; there were no synaesthetes experiencing both types), and 79 other-synaesthetes. Since this data collection was conducted on-line and thus these participants were not interviewed, in this subsample, other-synaesthetes were those

[26] Sample B completed the Multisense Consistency Test (MCT) to assess synaesthetic colour consistency, whereas Sample A (Study 3) completed either the MCT or the Synesthesia Battery (SB) due to technical problems during the data collection phase (see Footnote 13). However, the differences between the two tests are minimal and the MCT provides a comparable SB score. Similarly, the two subsamples performed slightly different versions of the Sussex´s Sequence-Spatial Synaesthesia Diagnostic Test: Sample B completed the standard version of the test and Sample A a version with minor changes to the wording with adjusted thresholds (see Footnote 14).

[27] Independent-samples *t*-tests comparing -colour and sequence consistency scores between participants who completed the ESSA first and then the personality questionnaires and participants who completed the assessments in the reverse order showed no differences between the two groups (both $t < 1.59$, both $p > .11$).

participants who failed both consistency tests, reported to have other types of synaesthesias that could not be assessed, and who additionally passed the score thresholds for the group defined by the Training sample analyses in Study 5 (see section 4.3.2.1). We applied the same restrictions regarding consistency criteria and types of synaesthesia excluded defined for Sample A; thus, the final Sample B was composed of 134 non-synaesthetes, 52 colour-synaesthetes, and 10 sequence-synaesthetes (see Tables 6 and 7 above).

The data collection processes for Studies 4 and 5 were done in collaboration with the University of Sussex (Brighton, UK) and followed the ethical guidelines laid down in the Helsinki Declaration. The project was further approved by The University of Edinburgh's Psychology Research Ethics Committee and the University of Sussex's Sciences & Technology Cross-Schools Research Ethics Committee. Participants were recruited via general and specialised social media on-line sites (e.g. University mailing lists, Facebook, or the 'Call for participants' website; but no synaesthesia groups or similar resources). Psychology undergraduates from The University of Edinburgh who participated in the study received study credits as compensation for their time (1 study credit/hour). The rest of the subjects did not receive any monetary compensation, but they were given the option to enter a prize draw of 2 x £100 and 10 x £50 vouchers (also open to The University of Edinburgh Psychology students). Informed consent was obtained from all subjects.

3.2.3 Data analyses (Samples A and B)

First, we calculated participants' scores on the different questionnaires according to the measurements' specifications (see section 3.2.1). Then, mean scores on each personality trait were calculated per group and sample[28] and submitted to Analyses of Variance (ANOVAs)

[28] The two samples were separately analysed because preliminary analyses (independent samples *t*-tests or non-parametric Mann-Whitney U) comparing the means of each personality trait between Samples A and B independently for each group (non-synaesthetes, colour-synaesthetes, and sequence-synaesthetes) revealed significant differences in several traits. In particular, differences between Samples A and B were found in BFI Conscientiousness ($t(163) = 2.86$, $p = .005$) and O-LIFE Unusual Experiences ($U(163) = 1{,}370$, $p = .003$) for non-synaesthetes; in BFI Openness to Experience ($t(163) = 2.55$, $p = .013$) for colour-synaesthetes; and in BFI Neuroticism ($t(163) = 2.19$, $p = .038$) and O-LIFE Cognitive Disorganisation ($U(163) = 32$, $p = .006$) for sequence-synaesthetes.

with ‘Group’ (non-synaesthetes, colour-synaesthetes, and sequence-synaesthetes) as the fixed factor[29]. Further post-hoc independent samples *t*-tests were carried out as appropriate following significant main effects. In case of violation of the assumption of normal distribution of the dependent variables (as assessed by Shapiro-Wilk), we used alternative, non-parametric Mann-Whitney U tests. In addition, Bonferroni corrections for multiple comparisons were applied adjusting the alpha threshold accordingly when needed.

We also examined whether synaesthetic strength was associated with personality traits. To that aim, we conducted Pearson correlations (or Spearman rank correlations for ordinal variables; e.g. Salkind, 2010; Schober et al., 2018) between those traits for which group differences were observed and three synaesthetic strength measures of interest: synaesthetic consistency scores, number of synaesthesia types, and overall degree of synaesthetic experience (as assessed with the ESSA screening interview; see Table 9 in section 3.3.2 for all the details about these measures’ sources). These analyses were separately conducted for colour- and sequence-synaesthetes, excluding non-synaesthetes (as, otherwise, any potential relationships detected could not have been distinguished from the group differences already observed). Bonferroni corrections were applied for multiple comparisons (adjusted $\alpha = .017$).

All analyses were conducted in Jamovi 0.9 (Jamovi Project, 2018) and SPSS 24 (IBM Corporation, 2016).

[29] To disregard any possible influences of handedness or level of education in Sample B (see Table 6), we run the same ANOVAs for each of the personality measures adding these variables as covariates (run separately). The analyses showed that ‘Group’ (non-synaesthetes, colour-synaesthetes, sequence-synaesthetes) and ‘Handedness’ interacted for BFI Agreeableness ($F(2, 189) = 1.12$, $p = .048$). However post-hoc independent *t*-tests (or Mann-Whitney U tests in case of violation of the assumption of normality, as assessed by Shapiro-Wilk) did not show any differences for any of the groups (all $p > .022$; Bonferroni-adjusted alpha = .017). Regarding ‘Level of education’, the analysis determined that this variable interacted with ‘Group’ for O-LIFE Cognitive Disorganisation ($F(4, 187) = 2.69$, $p = .033$). Post-hoc comparisons revealed that colour-synaesthetes with a Master level of education had significantly higher scores on the subscale than those colour-synaesthetes with a PhD ($U(10) = 0$, $p = .016$; Bonferroni-adjusted alpha = .0167). However, it is worth noticing that the subsample sizes for this analysis were very small (colour-synaesthetes with a Master = 9; with a PhD = 3) and it barely passed the significance threshold. Moreover, all the rest of the comparisons were not significant: all $p > .05$ (Bonferroni-adjusted alpha = .017).

3.3 Results

3.3.1 Group differences

Table 8 shows a summary of the different personality scores obtained per group and sample.

Table 8.
Mean scores by sample and group for the different personality subscales assessed in Study 4.

	Sample A			**Sample B**		
	Colour-synaesthetes *N* = 41	Sequence-synaesthetes *N* = 18	Non-synaesthetes *N* = 31	Colour-synaesthetes *N* = 52	Sequence-synaesthetes *N* = 10	Non-synaesthetes *N* = 134
BFI Extroversion	3.15 (.76)	3.04 (.72)	3.10 (.92)	2.93 (.82)	3.39 (.57)	3.02 (.85)
BFI Agreeableness	3.60 (.59)	3.83 (.62)	3.55 (.63)	3.69 (.67)	3.78 (.49)	3.77 (.60)
BFI Conscientiousness	3.55 (.55)	3.57 (.64)	3.72 (.62)	3.29 (.72)	3.47 (.65)	3.37 (.61)
BFI Neuroticism	3.22 (.60)	3.17 (.61)	3.14 (.73)	3.25 (.82)	3.71 (.67)	3.25 (.78)
BFI Openness to Experience	3.59 (.47)	3.96 (.54)	3.37 (.47)	3.33 (.50)	3.62 (.35)	3.55 (.62)
IRI Perspective Taking	18.6 (5.01)	20.9 (4.40)	18.6 (4.31)	18.5 (4.53)	19.6 (3.89)	18.8 (4.76)
IRI Fantasising	18.8 (4.32)	21.7 (4.27)	18.2 (5.25)	17.7 (5.98)	20.3 (4.74)	18.5 (5.44)
IRI Empathic Concern	20.3 (4.58)	22 (3.48)	20.5 (4.61)	19.4 (5.62)	21.4 (2.95)	20.2 (4.86)
IRI Personal Distress	14.5 (4.69)	13.8 (4.73)	12.3 (4.05)	13 (5.029	14.5 (4.90)	13.2 (4.28)
O-LIFE Unusual Experiences	.34 (.23)	.37 (.19)	.19 (.17)	.33 (.23)	.35 (.18)	.31 (.19)
O-LIFE Cognitive Disorganisation	.52 (.28)	.50 (.20)	.47 (.32)	.58 (.28)	.75 (.22)	.57 (.26)

Note: *N* = Sample size; Standard Deviations in parentheses; BFI = Big Five Inventory; IRI = Interpersonal Reactivity Index; O-LIFE = Oxford-Liverpool Inventory of Feelings and Experiences.

3.3.1.1 Sample A.

- Big Five Inventory (BFI): Significant group differences for the Openness to Experience subscale were found ($F(2, 87) = 8.52$, $p < .001$, $\eta_p^2 = .16$). Post-hoc comparisons with Bonferroni correction ($\alpha = .017$) indicated that sequence-synaesthetes (M = 3.96, SD = .52) experienced significantly higher rates than non-synaesthetes (M = 3.37, SD = .47; $t(47) = 4.02$, $p < .001$, $d = 1.19$) and colour-synaesthetes (M = 3.59, SD = .47; $t(57) = 2.65$, $p = .010$, $d = .75$) (Fig. 13A). Colour-synaesthetes and non-synaesthetes did not differ in their scores ($t(70) = 2$, $p = .05$). We also explored the four other subscales of the BFI with the alpha adjusted for multiple comparisons ($\alpha = .013$), but no group differences were observed for any of the traits (all $F(2, 87) < 1.29$, all $p > .28$).

- Interpersonal Reactivity Index (IRI): The analyses indicated group differences for the Fantasising subscale ($F(2, 87) = 3.59$, $p = .032$, $\eta_p^2 = .08$). Sequence-synaesthetes presented higher mean average scores compared to non-synaesthetes (M = 21.7, SD = 4.27 and M = 18.2, SD = 5.25, respectively), but the comparison did not survive corrections ($t(47) = 2.44$, $p = .018$; Bonferroni-adjusted $\alpha = .017$). The same pattern of responses was observed with respect to colour-synaesthetes (M = 18.8, SD = 4.32 $t(57) = 2.42$, $p = .019$, Bonferroni-adjusted $\alpha = .017$) (Fig. 13B). Colour-synaesthetes and non-synaesthetes showed no mean rate differences ($t(70) = .55$, $p = .59$). No differences were found either in the exploration of the rest of the IRI subscales (all $F(2, 87) < 2.17$, all $p > .12$; Bonferroni-adjusted $\alpha = .017$).

- Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE): The analyses revealed group differences for the Unusual Experiences subscale ($F(2, 87) = 6.54$, $p = .002$, $\eta_p^2 = .13$; Bonferroni-adjusted $\alpha = .025$). In this case, both sequence- and colour-synaesthetes (M = .37, SD = .19 and M = .34, SD = .23, respectively) had significantly higher rates than non-synaesthetes (M = 19, SD = 17; $t(47) = 3.44$, $p = .001$, $d = 1.02$ and $U(70) = 383$, $p = .004$, $d = .74$, respectively – Bonferroni-adjusted $\alpha = .017$) (Fig. 13C). No differences were found between sequence- and colour-synaesthetes ($t(57) = .48$, $p = .63$). There were no group differences either for the Cognitive Disorganisation subscale ($F(2, 87) = .27$, $p = .76$).[30]

[30] Since the BFI Agreeableness and O-LIFE Cognitive Disorganisation subscales violated the assumption of normality (as assessed by Shapiro-Wilk), we run additional non-parametric Kruskal-Wallis H tests. The analyses confirmed the no group differences for both variables ($\chi^2(2) = .45$, $p = .80$ and $\chi^2(2) = 2.39$, $p = .30$, respectively).

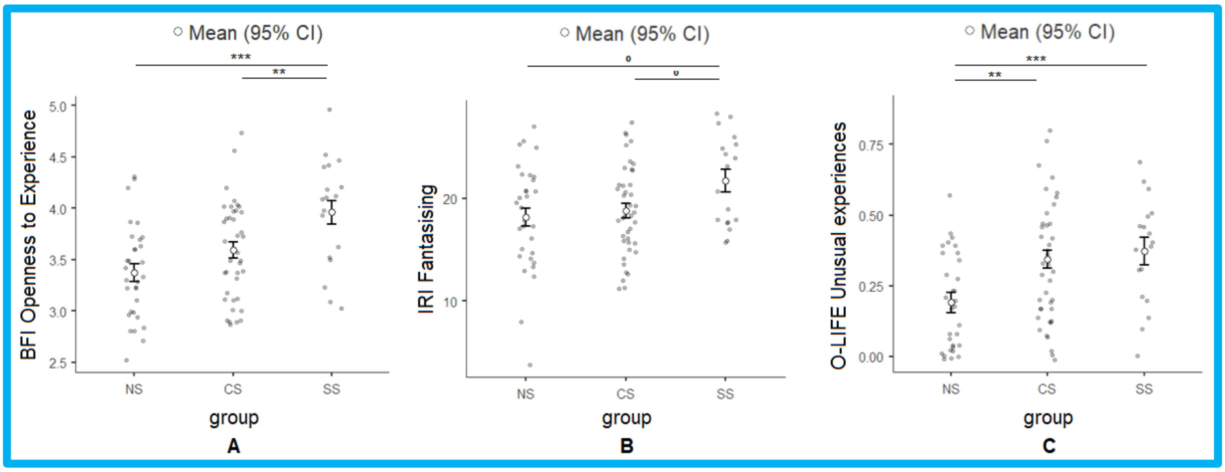


Figure 13. Mean scores on personality traits and their corresponding confidence intervals (error bars), for each group (non-synaesthetes – NS; colour-synaesthetes – CS; sequence-synaesthetes – SS) – Sample A of Study 4. Figure **A** Sequence-synaesthetes obtained significantly higher rates on the BFI (Big Five Inventory) Openness to Experience subscale compared to non-synaesthetes ($p < .001$) and colour-synaesthetes ($p = .010$) (Bonferroni-adjusted $\alpha = .017$). Figure **B** Group differences were also observed for the IRI (Interpersonal Reactivity Index) Fantasising subscale ($p = .032$), sequence-synaesthetes showing higher rates than the non-synaesthetes and colour-synaesthetes. However, post-hoc comparisons did not survive multiple comparisons corrections ($p = .019$ and $p = .017$, respectively). Figure **C** Finally, both colour- and sequence-synaesthetes experienced higher rates than non-synaesthetes on the O-LIFE (Oxford-Liverpool Inventory of Feelings and Experiences) subscale ($p = .001$ and $p = .004$, respectively).

3.3.1.2 Sample B.

- Big Five Inventory (BFI): Group differences only approached significance for the Openness to Experience subscale ($F(2, 193) = 2.89$, $p = .058$). The analysis exploring the other BFI subscales did not show either any group differences (all $F(2, 193) < 1.66$, all $p > .19$; Bonferroni-adjusted $\alpha = .013$).

- Interpersonal Reactivity Index (IRI): Contrary to our hypothesis and contradicting Sample A findings, the analyses revealed no group differences for the Fantasising subscale ($F(2, 193) = 1.05$, $p = .35$). No differences were found either in the examination of the rest of the IRI subscales (all $F(2, 193) < .84$, all $p > .43$; Bonferroni-adjusted $\alpha = .017$).

- Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE): Similarly, no group differences were observed in this sample for the Unusual Experiences or the Cognitive

Disorganisation subscales ($F(2, 193) = .41$, $p = .66$ and ($F(2, 193) = 2.19$, $p = .11$, respectively; Bonferroni-adjusted $\alpha = .025$).[31]

3.3.2 Synaesthetic strength differences (Sample A)

We only examined correlations between synaesthetic strength measures and those personality traits for which group differences were observed. The analyses revealed that the number of synaesthesia types (ESSA EN) and overall degree of synaesthetic experience (ESSA EH) self-reported were significantly and positively correlated with the O-LIFE Unusual Experience scores for the colour-synaesthetes subsample of Sample A ($r = .55$, $p < .001$ and $\rho = .43$, $p = .005$; Bonferroni-adjusted $\alpha = .017$) (Fig. 14A and 14B). That is, the greater the number of synaesthesia types and the higher the overall degree of synaesthetic experience, the higher the scores for the O-LIFE Unusual Experiences subscale. However, there was no relationship between this personality trait and synaesthetic -colour consistency scores ($r = .081$, $p = .61$).

Regarding the sequence-synaesthetes subsample, positive correlations between the overall degree of synaesthetic experience (ESSA EH) and the scores in IRI Fantasising and O-LIFE Unusual Experiences, but the relationships did not survive multiple comparison corrections (IRI Fantasising $\rho = .52$, $p = .026$ and O-LIFE Unusual Experiences $\rho = .53$, $p = .025$; Bonferroni-adjusted $\alpha = .017$). No associations were observed between these personality traits and the number of synaesthesia types (ESSA EN) or sequence consistency scores, or between BFI openness to experience and any of the measures evaluated (all $r/\rho < .33$, all $p > .19$; see Table 9 for the complete statistics).

[31] We run additional non-parametric Kruskal-Wallis H tests for those variables which violated the assumption of normality (i.e. BFI Agreeableness, BFI Neuroticism, IRI Perspective Taking, IRI Fantasising, IRI Empathic Concern, O-LIFE Unusual Experiences, and O-LIFE Cognitive Disorganisation). The analyses ratified the no group differences for all the variables (all $\chi^2(2) < 4.62$, all $p > .10$).

Table 9.
Correlations between the personality subscales and the synaesthetic strength measures analysed in Study 4, by group (Sample A).

Synaesthesia strength measure	*N*		BFI Openness to Experience	IRI Fantasising	O-LIFE Unusual Experiences
Colour-synaesthetes					
-Colour SB/MCT consistency score*	41	Pearson's *r* / *p*-value	-	-	.081 / .61
ESSA Extended Number (EN) score	41	Pearson's *r* / *p*-value	-	-	.55 / < .001º
ESSA Extended Highest (EH) score	41	Spearman's *ρ* / *p*-value	-	-	.43 / .005º
Sequence-synaesthetes					
Sequence SDT consistency score	18	Pearson's *r* / *p*-value	.10 / .68	.10 / .69	-.17 / .50
ESSA Extended Number (EN) score	18	Pearson's *r* / *p*-value	-.26 / .30	.04 / .88	.33 / .19
ESSA Extended Highest (EH) score	18	Spearman's *ρ* / *p*-value	.10 / .70	.52 / .026	.53 / .025

Note: *N* = Sample size; SB = Synesthesia Battery; MCT = Multisense Consistency Test; SDT = Sussex's Sequence-Spatial Synaesthesia Diagnostic Test; ESSA = Edinburgh Synaesthesia Screening Assessment (questionnaire version).[32]
* Scores of participants who completed the MCT were transformed to SB scores for homogenisation purposes.
º Significant relationships (Bonferroni-adjusted α = .017).

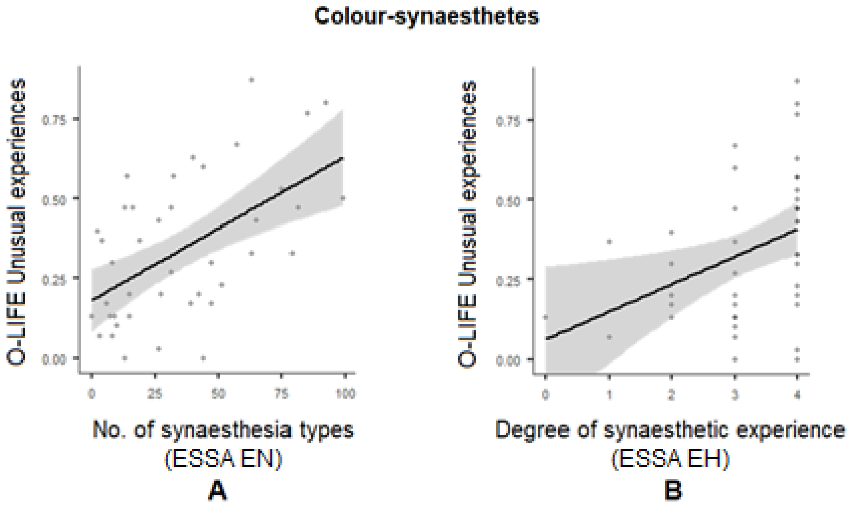


Figure 14. Correlations between personality traits and synaesthetic strength measures. The reported number of synaesthesia types (ESSA EN; Figure **A**) and overall degree of synaesthetic experience (ESSA EH; Figure **B**) were significantly and positively correlated with colour-synaesthetes scores on the O-LIFE (Oxford-Liverpool Inventory of Feelings and Experiences) Unusual Experiences subscale (r = .55, p < .001 and ρ = .43, p = .005; Bonferroni-adjusted α = .017).

3.4 Discussion

The present study directly investigated whether different types of synaesthetes present distinct personality profiles. Previous evidence has shown that synaesthetes have a distinct personality profile compared to non-synaesthetes, but there are inconsistencies in the

[32] Given that not all participants completed the interview version of the ESSA (see Footnote 24), questionnaire responses were used.

literature with respect to the personality traits that differ. Most studies have focused on grapheme-colour synaesthetes, ignoring other types of synaesthesia. However, some initial evidence suggests that sequence-synaesthetes might not experience the same personality profile as other synaesthetes (Ward et al., 2018a). In order to address this, we compared matched groups of colour-synaesthetes (e.g. letters-colours), sequence-synaesthetes (e.g. number- or calendar-forms), and non-synaesthete controls on the Big Five personality traits: Extraversion, Agreeableness, Neuroticism, Conscientiousness, and Openness to Experience (as assessed by the Big Five Inventory or BFI). Taking into account previous findings in the literature, participants also completed specific measures on empathy (Interpersonal Reactivity Index or IRI; Perspective Taking, Fantasising, Empathic Concern, and Personal Distress subscales) and schizotypy (Oxford-Liverpool Inventory of Feelings and Experiences or O-LIFE; Unusual Experiences and Cognitive Disorganisation subscales).

We replicated the findings observed in previous studies that synaesthetes, compared to controls, experience higher rates of BFI Openness to Experience (Banissy et al., 2013a; Chun & Hupé, 2016; Rouw & Scholte, 2016); IRI Fantasising, a dimension of empathy (Banissy et al., 2013a; Chun & Hupé, 2016; Rader & Tellegen, 1987; Rouw & Scholte, 2016); and O-LIFE Unusual Experiences (positive schizotypy; Banissy et al., 2012; Janik McErlean & Banissy, 2016). However, contrary to existing literature, some of these differences were only observed for sequence-synaesthetes. Specifically, this group of synaesthetes showed significantly higher rates of BFI Openness to Experience than non-synaesthetes *and* colour-synaesthetes. In addition, it should be noted that colour- and non-synaesthetes did not differ in their scores. The same pattern was observed in relation to IRI Fantasising (i.e. higher scores for sequence-synaesthetes compared to both colour- *and* non-synaesthetes), but the group comparisons did not survive corrections. On the other hand, both colour- and sequence-synaesthetes showed significantly higher rates for O-LIFE Unusual Experiences than controls, and there were no differences between the two groups of synaesthetes for this trait. These results are important in relation to previous findings in the field. The specific differences observed for sequence-synaesthetes and, crucially, the lack of differences between colour-synaesthetes and controls might require a reinterpretation of the results published so far. If unknown (or unacknowledged) sequence-synaesthetes were present in the samples, it could be possible that the personality effects observed ought to be attributed to the sequence-synaesthete difference detected here rather than to any grapheme-colour or general synaesthetic particularities.

The findings of the present study suggest that all synaesthetes might not share the same personality profile – or, at least, that sequence- and colour-synaesthetes show differences in this respect. This seems to be in accordance with other investigations' observations that sequence-synaesthetes might be a particular group of synaesthetes. These individuals have showed cognitive advantages in areas such as memory (Brang et al., 2010; Simner et al., 2009b), time manipulation (Mann et al., 2009), mental rotation and visual imagery (Brang et al., 2010; Lunke & Meier, 2018; Simner et al., 2009b), spatial processing (Hale et al., 2014), or visual perception (Ward et al., 2017a). Furthermore, several studies have consistently reported higher rates of (self-reported) visual imagery for this subgroup of synaesthetes (Havlik et al., 2015; Price, 2009; Spiller & Jansari, 2008; Spiller et al, 2015; Ward et al., 2018a). This is especially interesting considering that the parietal cortex, which has been shown to be play a role in cognitive control during visual perception and imagery processes in non-synaesthetes (Ganis, Thompson, & Kosslyn, 2004), has also shown structural and functional differences in synaesthetes (see Dojat, Pizzagalli, & Hupé, 2019; Hupé & Dojat, 2015; and Rouw et al., 2011 for reviews). Moreover, this brain area has also been linked to the Openness to Experience trait (Kennis, Rademaker, & Geuze, 2013; Ricelli, Toschi, Nigro, Terracciano, & Passamonti, 2017).

However, it should be noted that not all these findings were replicated in a second sample analysed (Sample B). Although sequence-synaesthetes also showed higher rates of BFI Openness to Experience and IRI Fantasising compared to colour-synaesthetes and controls, group differences only approached significance for BFI Openness to Experience.

Similarly, both colour- and sequence-synaesthetes had higher rates of O-LIFE Unusual Experiences than non-synaesthetes, but the analyses did not reveal any group differences. There might be several reasons that explain this lack of effects. In the first place, while both samples included a similar number of synaesthetes (Sample A *N* = 59; Sample B *N* = 62), the distribution of types of synaesthetes was not equivalent: in Sample A there were 41 colour-synaesthetes and 18 sequence-synaesthetes, whereas in Sample B there were, respectively, 52 and 10 synaesthetes of each type. In addition, non-synaesthetes were unbalanced in the two samples: Sample A *N* = 31 and Sample B *N* = 134. Given that the key findings in Sample A involved differences of sequence-synaesthetes with respect to the other groups, it could be possible that this group was underpowered in Sample B to produce any meaningful effects (the trend for Openness to Experience conceivably supporting this hypothesis).

Moreover, preliminary analyses comparing the two samples showed significant differences, amongst others, for the scores for BFI Openness to Experience and O-LIFE Unusual Experiences (assessed independently for each group; see Footnote 28). Thus, sample size differences aside, this indicates that the participants of each sample had different response patterns. Although both samples were naïve to the purposes of the questionnaire and completed the ESSA on-line before doing any synaesthetic consistency tests, Sample A came afterwards to the Lab to complete additional tasks for other studies. It is possible that this condition might have affected the way participants responded to the different assessments. Several studies have reported that self-referred and volunteer participants score generally higher in certain personality traits, including Openness to Experience (e.g. Brüggen & Dholakia, 2010; Dollinger & Leong, 1993; Lönnqvist et al., 2007; Marcus & Schütz, 2005). Neither Sample A nor Sample B participants were self-referred, but Sample A subjects clearly scored overall higher in all personality traits (and in the synaesthesia screening questionnaire, see section 4.4). Thus, it might be argued that the fact that Sample A participants accepted to do the on-line questionnaires and tasks and to come to the lab afterwards (i.e. accepted a higher implication in the study and time commitment), made them inherently more motivated participants than Sample B subjects, who only had to respond to the on-line tasks.

In fact, contextual considerations and variation in study designs, albeit small, are some of the reasons considered to contribute to replicability issues (e.g. Tackett et al., 2017). During recent years, the field of psychological science has undergone a process of self-criticism regarding its scientific practices and the derived credibility of its findings (e.g. Lilienfeld & Waldman, in 2017; Pasher & Wagenmakers, 2012). The publication of high-profile replication failures (e.g. Donnellan, Lucas, & Cesario, 2015; Klein et al., 2014; Open Science Collaboration, 2015) has led to conversations of questionable research practices in the area. Some of the most relevant ones include the use of multiple alternative variables and statistical analyses, adaptive stopping rules when collecting data, exclusion or refinement of study conditions or sample strata, HARKing (i.e. hypothesising after the results are known), or p-hacking (i.e. reporting results which confirm hypothesis while ignoring disconfirming results) (see e.g. LeBel, 2015; Stevens, 2017; or Tackett et al., 2017 for in-depth literature on the topic). Besides these malpractices, this new interest into replicability and reproducibility failure has highlighted the fact that the many decisions which researchers take through the scientific process, from preferred sampling methods to the specific choices regarding experimental designs or data analyses, can all lead to different biases that ultimately add to noise in the data.

Moreover, synaesthesia research, like other areas such as clinical or comparative psychology, present a series of particularities that make them especially susceptible to replication issues (e.g. Stevens, 2017; Tackett et al., 2017). These fields are characterised for typically studying small sample sizes, which are associated with low statistical power. Having few individuals also increases repeated testing and exploratory data analyses in an aim to extract as much data as possible, as well as it encourages publishing small and non-significant effects. Researchers do not only deal with small samples, but they also might be constrained in terms of which individuals they can find and how they can find them. This results in high individual variability within subjects from study to study which might cause poor replication and fluctuations in effect sizes. As observed in the present study (see Fig. 11) variability within individuals is particularly enhanced in the synaesthetic field. This is explained by the small sample factor described but also because there is not a general agreed consensus on what synaesthesia is and is not and because there are multiple ways to assess it (see section 4.1 for an extended discussion). Therefore, the comparison of synaesthetes across studies might be particularly complicated. This overall data 'messiness' makes it difficult to operationalise what qualifies as a replication success or failure.

Tackett and colleagues (2017) propose a series of general recommendations to deal with the presented issues in clinical psychology that can be broadly applied in synaesthesia research. For example, they suggest reducing the number of questionable research practices by staying educated on concerns such as p-hacking and by pre-registering study hypotheses and consider sharing open datasets. They also advocate for defining replicability within the different specific field, considering its particularities. Lastly, in order to enhance power and robustness, they recommend the systematic conduction of power analyses as well as the incorporation of multiple measurements of key constructs, which should be harmonised across labs as much as possible. The last point is of great importance to the synaesthetic field due to the divergences in the definition of the condition and its measurements. Although synaesthetic consistency tests have become the 'gold-standard' of synaesthesia assessment, they are not exempt from their own problems, one of the most important being the fact that only a few synaesthesia types can be assessed through such tests (see section 4.1 for further details). Additional measures such as behavioural tests like the synaesthetic Stroop task or the embedded shapes task (see section 2.1.3) could be incorporated. But, perhaps, the key question in synaesthetic measurement is whether synaesthesia can in fact be measured (or to what extent) and, consequently, whether we should rely in the first instance in phenomenological approaches or more qualitative methodologies such as the self-report questionnaires and interviews proposed in this project.

Nevertheless, in order to try to understand further the influence of individual variability in the present study, we decided to examine here in an exploratory approach both-synaesthetes (i.e. individuals who experienced both types) and other-synaesthetes (i.e. individuals who failed both -colour and sequence-space consistency tests but reported having other types of synaesthesias that could not be assessed). To do this, we compared these groups' scores on the different personality traits to the main groups. Specifically, we submitted both/other-synaesthetes mean rates on each personality trait to Analyses of Variance (ANOVAs) with 'Group' (non-synaesthetes, colour-synaesthetes, sequence-synaesthetes, both/other-synaesthetes) as the fixed factor, separately for each sample (there were no both-synaesthetes in Sample B, therefore, these analyses were only performed for Sample A).

The analyses showed significant group differences in BFI Opennes to Experience for both-synaesthetes (Sample A) ($F(3, 113) = 6.72$, $p < .001$, $\eta_p^2 = .15$). In particular, both-synaesthetes experienced significantly higher rates than non-synaesthetes (M = 3.37, SD = .47; $t(56) = 3.36$, $p = .001$, $d = .89$; Fig. 15A). No differences were observed between both- and colour- (M = 3.59, SD = .47; $t(66) = 1.92$, $p = .060$) or sequence-synaesthetes (M = 3.96, SD = .54; $t(43) = .69$, $p = .50$). Group differences were also observed for the O-LIFE Unusual Experiences subscale ($F(3, 113) = 8.26$, $p < .001$, $\eta_p^2 = .18$; Bonferroni-adjusted $\alpha = .025$). Both-synaesthetes (M = .46, SD = .23) presented significantly higher rates than non-synaesthetes (M = .19, SD = .17; $t(56) = 5.10$, $p < .001$, $d = 1.34$; Fig. 15B). The same pattern was observed with respect to colour-synaesthetes (M = .34, SD = .23), but the comparison did not survive corrections ($t(66) = 2.04$, $p = .046$ Bonferroni-adjusted $\alpha = .017$). Both- and sequence-synaesthetes (M = .37, SD = .19) did not differ in their scores ($t(43) = 1.31$, $p = .20$). No other group differences were found for the rest of the traits (all $p > .027$; Bonferroni-adjusted $\alpha = .013$).

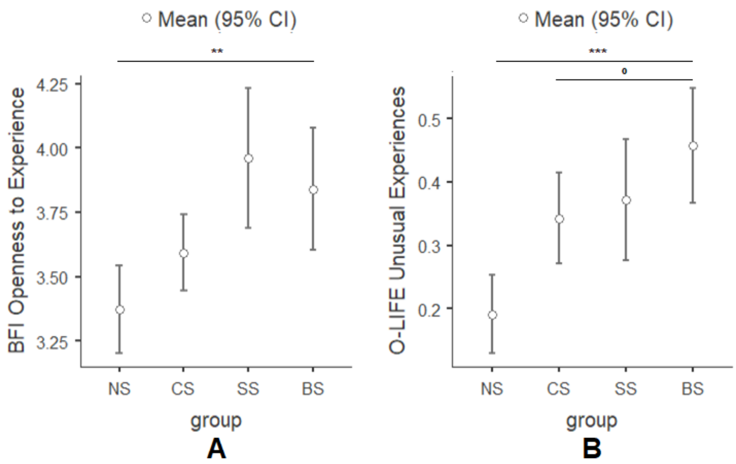


Figure 15. Mean scores on personality traits and their corresponding confidence intervals (error bars), for each group (non-synaesthetes – NS; colour-synaesthetes – CS; sequence-synaesthetes – SS; both-synaesthetes – BS) – Sample A of Study 4. Figure **A** Both-synaesthetes obtained significantly higher rates on the BFI (Big Five Inventory) Openness to Experience subscale compared to non-synaesthetes (p = .001; Bonferroni-adjusted α = .017). Figure **B** Both-synaesthetes also showed significantly higher rates than non-synaesthetes (p < .001) in the Unusual Experiences subscale of the O-LIFE (Oxford-Liverpool Inventory of Feelings and Experiences). A similar trend was observed with respect to colour-synaesthetes (p = .046), but the comparison did not survive corrections.

Regarding other-synaesthetes, the analyses for Sample A revealed group differences for BFI Openness to Experience (F(3, 125) = 5.65, p = .001, η_p^2 = .12). Other-synaesthetes (M = 3.70, SD = .55) experienced significantly higher rates than non-synaesthetes (M = 3.37, SD = .47; t(68) = 2.65, p = .010, d = .64; Fig. 16A). No differences were observed between other-synaesthetes and colour-synaesthetes (M = 3.59, SD = .47; t(78) = .92, p = .36) or sequence-synaesthetes (M = 3.96, SD = .54; t(55) = 1.70, p = .095). Group differences were also found for the O-LIFE Unusual Experiences subscale (F(3, 125) = 4.75, p = .004, η_p^2 = .10; Bonferroni-adjusted α = .025). Similarly, other-synaesthetes (M = .37, SD = .27) presented significantly higher rates than non-synaesthetes (M = .19, SD = .17; U(68) = 348, p = .002, d = .81; Fig. 16B), but no differences were observed between other-synaesthetes and colour- (M = .34, SD = .23; t(78) = .45, p = .66) or sequence-synaesthetes (M = .37, SD = .19; t(55) = .09, p = .93).

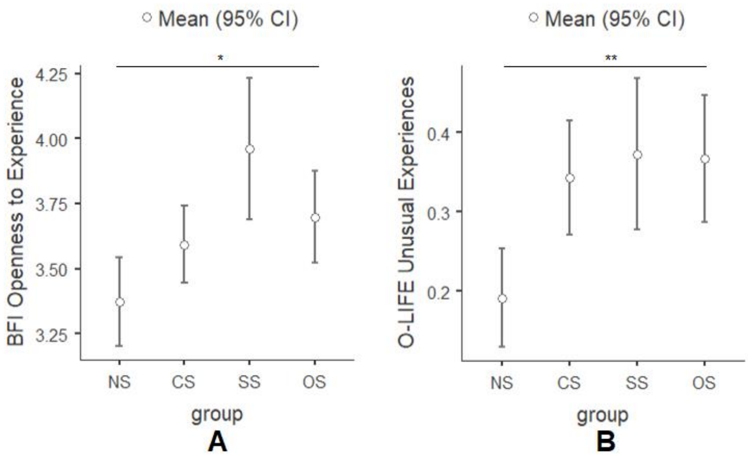


Figure 16. Mean scores on personality traits and their corresponding confidence intervals (error bars), for each group (non-synaesthetes – NS; colour-synaesthetes – CS; sequence-synaesthetes – SS; other-synaesthetes – OS) – Sample A of Study 4. Figure **A** Other-synaesthetes obtained significantly higher rates on the BFI (Big Five Inventory) Openness to Experience subscale compared to non-synaesthetes (p = .010; Bonferroni-adjusted α = .017). Figure **B** Other-synaesthetes also showed significantly higher rates than non-synaesthetes (p = .002) in the Unusual Experiences subscale of the O-LIFE (Oxford-Liverpool Inventory of Feelings and Experiences).

Group differences were also observed in O-LIFE Unusual Experiences for Sample B (F(3, 271) = 10.3, p < .001, η_p^2 = .10; Bonferroni-adjusted α = .025). In this case, other-synaesthetes (M = .46, SD = .20) presented significantly higher rates than non-synaesthetes (M = .30, SD = .19; t(211) = 5.61, p < .001, d = .80) *and* colour-synaesthetes (M = .33, SD = .23; t(129) = 3.51, p < .001, d = .63) (Fig. 17). Other- and sequence-synaesthetes (M = .35, SD = .18) did not differ in their scores (t(87) = 1.69, p = .094). The other-synaesthetes analyses for Sample A and Sample B did not show any other group differences for the rest of the traits (all p > .063) (the complete analyses for both samples and types of synaesthetes can be consulted in Appendix F).

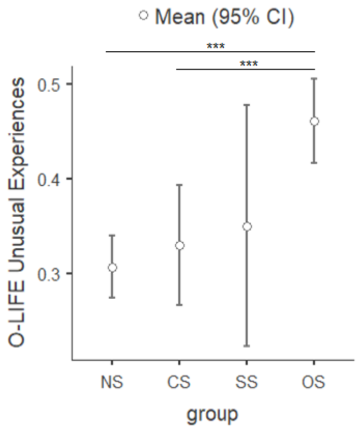


Figure 17. Mean scores on the O-LIFE (Oxford-Liverpool Inventory of Feelings and Experiences) Unusual Experience subscale and their corresponding confidence intervals (error bars), for each group (non-synaesthetes – NS; colour-synaesthetes – CS; sequence-synaesthetes – SS; other-synaesthetes – OS) – Sample B of Study 4. Other-synaesthetes obtained significantly higher rates on this personality trait compared to non-synaesthetes (p < .001) and colour-synaesthetes (p < .001) (Bonferroni-adjusted α = .017).

Therefore, both- and other-synaesthetes showed similar personality trait differences compared to non-synaesthetes as those observed for sequence-synaesthetes in the main analyses (i.e. higher rates for BFI Openness to Experience and O-LIFE Unusual Experiences). Importantly, these groups did not show differences with respect to sequence-synaesthetes. These results might suggest thus that perhaps colour-synaesthetes are the odd group with a particular personality profile.

Besides different types of synaesthesias, synaesthetes can also have greater or fewer synaesthesia types, present different scores of synaesthetic consistency for a specific type of synaesthesia, or report different degrees of overall synaesthetic experience. A secondary aim of the present study was to examine whether these individual differences in synaesthetic strength modulated the ratings in those personality traits that differentially characterise synaesthetes from controls. The results showed that colour-synaesthetes showed significant positive correlations between O-LIFE Unusual Experiences rates (i.e. positive schizotypy) and the number of synaesthesia types (ESSA EN) and the overall degree of synaesthetic experience reported (ESSA EH) – but not with synaesthetic -colour consistency scores. This means that the higher the number of synaesthesia types (ESSA EN) / the higher the overall degree of synaesthetic experience (ESSA EH), the higher the rates for the O-LIFE Unusual

Experiences subscale. Regarding sequence-synaesthetes, this group showed positive correlations between the overall degree of synaesthetic experience (ESSA EH) and the rates of IRI Fantasising and O-LIFE Unusual Experiences, but the relationships did not survive multiple comparison corrections.

These findings are in partial agreement with previous literature. On the one hand, like Hossain et al. (2018), we did not find any relationships between personality traits and (-colour or sequence) synaesthetic consistency scores. On the other hand, in contrast to Rouw and Scholte (2016), we did not observe positive associations between the number of synaesthesia types (ESSA EN) and rates of IRI Fantasising or BFI Openness to experience. Moreover, our results showed significant positive correlations between rates of O_LIFE Unusual Experiences (i.e. positive schizotypy) and number of synaesthesia types (ESSA EN) / overall degree of synaesthetic experience (ESSA EH). Whilst all these findings should be taken with caution as the analyses were conducted on separate colour- and sequence-synaesthetes subsamples and, therefore, the sample sizes were small, the observed dissociation between qualitative and quantitative effects of synaesthesia on personality highlight the need to consider synaesthetic strength as a relevant factor or variable of interest in future studies.

In sum, the present study results ratified previously reported synaesthetic differences in BFI Openness to Experience, IRI Fantasising, and O-LIFE Unusual Experiences (i.e. positive schizotypy) in relation to non-synaesthetes. However, we have observed a critical difference: some of these differences were only for a specific group of synaesthetes, namely sequence-synaesthetes. Sequence-synaesthetes had significantly higher rates of BFI Openness to Experience than non-synaesthetes and, crucially, *than* colour-synaesthetes. The same pattern of responses was observed for the empathy trait of IRI Fantasising, but group comparisons did not survive multiple corrections. Furthermore, it should be noted that exploratory analyses comparing additional groups of synaesthetes (both-synaesthetes, i.e. individuals with -colour and sequence-synaesthesias; and other-synaesthetes, i.e. subjects with other types of synaesthesias) revealed that these groups showed a similar personality profile to that of sequence-synaesthetes, suggesting perhaps the singularity of colour-synaesthetes On another note, stronger degrees of synaesthetic strength (measured as the number of synaesthesia types, ESSA EN, and as the overall degree of synaesthetic experience reported, ESSA EH) were related to higher rates of O-LIFE Unusual Experiences for colour-synaesthetes. Similar positive trends were observed for sequence-synaesthetes with respect to ESSA-EH and IRI Fantasising and O-LIFE scores, but the relationships did not survive multiple comparison corrections (maybe because this smaller subsample was underpowered; 18 sequence- vs. 41 colour-synaesthetes). Therefore, the present study represents the first direct evidence that not all synaesthetes seem to have the same personality profile and it has highlighted the importance of considering individual differences in synaesthetes both in terms of types and strength. However, no differences between synaesthetes and non-synaesthetes or between the two types of synaesthetes were observed in a second sample assessed. Although we have discussed several possible limitations that could have affected this sample, further studies are required to corroborate these inter- and intra-synaesthetic differences in personality in new samples paying special attention on subject recruitment and assessment administration methods. In addition, it would be interesting to explore differences between other types of synaesthetes and address more systematically the synaesthetic strength variable in future investigations.

4. Chapter IV: Development and Validation of the Edinburgh Synaesthesia Screening Assessment – ESSA (Study 5)

4.1 Introduction

Knowing the synaesthetic status and characteristics of participants are essential requirements to investigate any aspect of synaesthesia or of synaesthetes. First, there is a need to clearly establish who is a synaesthete and who is not. Most studies will (more or less thoroughly) ensure that their experimental group consists of people who have synaesthesia. However, not the same degree of meticulosity is applied with respect to the control group. Only a minority of studies assess controls with the same tests and questionnaires completed by synaesthetes or with the same detail. In fact, is not unusual that the assessment of controls is limited to simple “Do you have synaesthesia?” or “Do you have [this particular] type of synaesthesia?” questions. This poses a problem for individuals who are unaware of their synaesthetic experiences and/or do not consider them ‘strong’ or ‘unusual’ enough to report them.

In addition, differences within synaesthetes are often neglected. Synaesthetes can present different degrees of synaesthetic strength, frequencies of experience, locations (associators vs. projectors), or, critically, types of synaesthesias. Growing evidence is showing that these differences might be more relevant than previously thought, in some cases being fundamental (e.g. Dixon & Smilek, 2005; Hale et al., 2014; Havlik et al., 2015; Hossain et al., 2018; Jonas & Price, 2014; Lunke & Meier, 2018; Spiller et al., 2015; Ward et al., 2007; Ward et al., 2017a). Most studies will evaluate participants on the target synaesthesia(s) which want to be investigated – or, if this is not an important factor, on the most common types. However, questions regarding additional types of synaesthesia are not always included and, if they are, they do not tend to be exhaustive and are limited to a few extra types. This is both a problem in terms of considering the full synaesthetic scope of synaesthete participants and of detecting less prevalent types of synaesthetes. Similarly, the collection of additional synaesthetic characteristics is not common practice (except for the associator/projector variable).

Several synaesthesia assessment and screening tools have been developed over the years to address these problematics with varying approaches (see Johnson et al., 2013 for a review). Synaesthetic consistency tests are considered the ‘gold-standard’ of synaesthetic measurement and their use is widespread in contemporary synaesthetic research. They rely on the premise of synaesthetic constancy or the fact that inducer-concurrent associations tend to be highly stable over time and for the person, which is considered one of the main defining criteria of synaesthesia (e.g. Ward, 2013). Initially, these assessments took the form of simple test-retests and consisted of recording the individual’s synaesthetic associations and comparing the person’s responses again after a determined period of time (e.g. Dresslar, 1903; Ginsberg, 1923).

Baron-Cohen and colleagues (1987; 1993; 1996) updated and improved this format with the Test of Genuineness (TOG) for grapheme-colour synaesthesia, which introduced lists of specific word and letter triggers. Importantly, they also developed an independent-raters scoring system and compared responses of synaesthetes and matched controls. With the aim to work towards a more standardised measure, years later Asher, Aitken, Farooqi, Kurmani, and Baron-Cohen (2006) created the Revised Test of Genuineness (TOG-R), with a new list of triggers which included words, letters, and also sounds. Importantly, the TOG-R switched verbal descriptions for a chart with 238 numbered coloured swatches from which participants had to choose and the scoring system was also reviewed: different points were given depending on the proximity of the colour swatches chosen in the test and the retest, closer proximities obtaining higher points and reflecting higher consistency.

More recently, the tests of genuineness for grapheme- and sound-colour synaesthesias evolved to computerised versions with much more refined and powerful colour palettes (over 16 million colour bitmaps) and more sophisticated algorithms to calculate colour spectral distances (e.g. Eagleman et al., 2007; Menouti, Akiva-Kabiri, Banissy, & Stewart, 2015; Simner & Ludwig, 2012; Simner et al., 2009a; MULTISENSE Research Project, 2019). It is also worth mentioning that some of these new tests (e.g. Eagleman, 2007; Menouti et al., 2015; MULTISENSE Research Project, 2019) are completed as single testing sessions (i.e. the different triggers are shown over repeated trials instead of time points), facilitating diagnostic. Noticeably, Eagleman's Synesthesia Battery (2007) has become one of the most widely used paradigms in synaesthesia research and its reliability to detect synaesthesia has been firmly validated (e.g. Carmichael et al., 2015*;* Eagleman et al., 2007; Rothen et al., 2013). In addition, consistency tests based on similar principles and procedures have also been developed for other synaesthetic triggers such as sequence-spatial synaesthesias (e.g. Brang et al., 2013a; Cytowic, 2002; Eagleman, 2009; Rothen, Jünemann, Mealor, Burckhardt, & Ward, 2016; Sagiv et al., 2006b; Ward et al., 2018a – see section 2.2.2.1.1 for details), ordinal-linguistic personifications (e.g. Smilek et al., 2007b; MULTISENSE Research Project, 2019), lexical-gustatory synaesthesia (e.g. Ipser, Ward, & Simner, 2019; Ward & Simner, 2003), or – saving some methodological distances – mirror-touch synaesthesia (e.g. Baron-Cohen et al., 2016; Holle, Banissy, Wright, Bowling, & Ward, 2011; Ward, Schnakenberg, & Banissy, 2018b).

In sum, synaesthetic consistency tests have proven to be a robust and objective method to distinguish between synaesthetes and non-synaesthetes (i.e. mean consistency scores for the two populations that do not typically overlap – Asher et al., 2006; Baron-Cohen et al., 1993; 1996; Ward, Simner, & Auyeung, 2005). Moreover, the format can be adapted to different trigger stimuli and the computerised versions are particularly easy to administer and made the test more accessible for diverse audiences (e.g. children or verbally-impaired populations). However, they also present some limitations. In the first place, they rely on the assumption of synaesthetic consistency, but some studies are starting to question the stability of synaesthetic associations across the lifespan (e.g. Meier et al., 2014; Simner et al., 2009a; Simner et al., 2017) or in other circumstances such as mood changes (e.g. Kay, Carmichael, Ruffell, & Simner, 2015).

Secondly, although the fact that controls perform poorer in these tests despite giving them 'advantages' (e.g. warnings about retest, shorter retest intervals) is considered a further prove of the authenticity of synaesthesia captured, false positive diagnoses are rare but possible. Deliberate memory or cheating strategies (e.g. by taking note of the associations) could be employed, especially when remote computerised administrations without the presence of the researcher take place. However, the use of complementary interviews or self-report questionnaires has been observed to help to reduce this risk of false positives (e.g. Simner et al., 2006; Ward et al., 2006). Fatigue and frustration are also disadvantages which need to be taken into account and which can cause false negatives. The introduction of large detailed colour charts allowed synaesthetes to finely select their perceived colours, but sometimes this comes at the cost of choice overwhelm (Asher et al., 2006). In addition, when many triggers are presented, the test becomes an arduous task to some participants (especially controls, for whom associations are meaningless). In order to address this problem, some studies have used versions with fewer items, observing that results were diagnostically accurate as well (e.g. Asher et al., 2006; Baron-Cohen et al., 1987; Simner et al., 2006).

However, the greatest limitation of synaesthetic consistency tests is that only a reduced number of synaesthesia types can be assessed. While similar measures could be potentially adapted to many synaesthesia types, the idea to test an individual for all possible types of synaesthesias would not be neither realistic nor feasible. There are at least 73 different types of synaesthesias known to date (Day, 2019), so completing that many consistency tests would be extremely time consuming and extenuating for both participants and researchers. Self-report interviews and questionnaires can be good methodological tools to screen both synaesthetes and non-synaesthetes and address synaesthetic variability at the same time, exploring different types of synaesthesia in a much more time- and cost-efficient way. That is, they can be used as a first step to narrow down the decision of which consistency tests will be administered to ultimately verify specific synaesthesia types with consistency measures.

Several studies have designed self-report questionnaires/interviews for screening purposes (e.g. Banissy & Ward, 2007; Banissy et al., 2009; Baron-Cohen et al., 1993; Chun

& Hupé, 2013; Kusnir & Thut, 2012; Rothen & Meier, 2010; Rouw & Scholte, 2007; 2016; Simner et al., 2009a; Simner et al., 2006). However, these measures were adapted to each study's investigation needs. This means that a variety of methods have been employed in terms of types and number of synaesthesias included (usually not exhaustive and focused on the aims of the particular study), of question types (e.g. "How much do you identify yourself with…", "Do you experience…"), of ratings options (e.g. "Yes / No", "On a scale from 1 to 5", "Always / Sometimes / Never"), of administration modalities (interview/questionnaire, provision of examples needed or not, etc.), and of classification criteria and/or scorings systems applied (if they used any, and more or less detailed).

This lack of systematicity does not only imply different results obtained from study to study, but also the impossibility to ascertain the comparability of populations. In other words, a person classified as synaesthete in one study could fall into the control group in another. This is partly caused by the different specifications and criteria defined, but primarily due to the fact that different synaesthesia types are included. In order to address these problems and work towards a standardised synaesthesia screening measurement, here we present the development and validation of the Edinburgh Synaesthesia Screening Assessment (ESSA), a self-report screening tool for synaesthesia designed to be broadly used in synaesthesia research regardless of the specific objectives of a study. For that reason, the measure covers an exhaustive range of synaesthesia types at the same time that it aims to be adequate to assess both synaesthetes and non-synaesthetes. To evaluate the predictive power of the ESSA, we analysed the optimal thresholds to distinguish between synaesthetes and non-synaesthetes against participants' grapheme-colour and sequence-spatial consistency scores. Lastly, we conducted internal and external (different sample) validation analyses.

4.2 Methods

4.2.1 Development and characteristics of the ESSA

The Edinburgh Synaesthesia Screening Assessment (ESSA), which can be either administered as an interview or as a questionnaire, originated from the need to determine the synaesthetic status of the participants of our behavioural and personality studies. One of our main goals was to develop a screening tool that could be administered to both synaesthetes and non-synaesthetes. To that end, we decided that a good approach was to focus on the experiential aspect of synaesthesia. That is, we designed a series of statements that described the experience of each particular synaesthesia type, in a language as simple and clear as possible and omitting the word synaesthesia or other jargon related to it (e.g. inducer or concurrent). We also tried to use descriptions and words that were inclusive of the different possible ways to experience synaesthesia (e.g. “when I see, hear, or think” or “perceive” instead of “see”, to target better both projector and associator experiences). For instance, the statement regarding letters-colours synaesthesia was formulated as: “I perceive different colours when I see, hear, or think about the letters of the alphabet”.

In addition, we decided to use a rating question and response scale that could reflect different degrees of experience – including the no-experience for non-synaesthetes. In the initial or pilot version of the screening tool (Studies 1 and 2; see section 2.2.2.1.1 and Appendix A) we explored different characteristics of the synaesthetic experience (i.e. frequency, constancy, location, and stability) and, adapting from Banissy et al. (2009) and Kusnir and Thut (2012), we took the frequency aspect (i.e. “How often do you experience this?”; answer options: ‘Always’, ‘Sometimes’, ‘Never’) as the criterium to classify participants. People who reported having any of the presented statements ‘Always’ were considered synaesthetes, whereas people who chose the option ‘Never’ for all the statements were classified as non-synaesthetes. People who only reported having synaesthetic experiences ‘Sometimes’ were cautiously grouped as ‘weak’ synaesthetes. In the revised version of the screening tool, used in Study 3 (see section 2.4.2.1.1 and Appendix D), we changed the classifying criterium as we considered that focusing on the frequency aspect could be limiting. Therefore, we decided to change the rating question and response scale to a broader formulation that were as neutral as possible and not linked to any specific characteristic of the synaesthetic experience. The new rating question was formulated as “How much does this experience applies to you?” and offered a 5-point Likert response scale: ‘Not at all’, ‘A little bit’, ‘Moderately’, ‘Quite a lot’, ‘Completely’.

Our second aim was to have a questionnaire that was as exhaustive as possible regarding the types of synaesthesias explored and that did not focus only on the most known and studied ones (e.g. -colour synaesthesias, sequence forms, mirror-touch synaesthesia, etc.) so that the number of undetected synaesthetes could be reduced. To decide which synaesthesias to include, we took the ESSA pilot version as a starting point, removing those statements (i.e. synaesthesia types) which we had identified were not working – because they were unclear, they promoted ambiguous answers, or they were just not representative of synaesthetic associations. We then examined several publications on the prevalence of synaesthesia types (e.g. Johnson et al., 2013; Simner et al., 2006; Ward, 2013). We also took Dr. Day’s web register (2019) as a reference since it lists prevalence data on synaesthesia types based on self-declared records of 1,143 synaesthetes. Following next, we organised the selected synaesthesia types by concurrents following the findings of a large study conducted by Novich et al. (2011). The authors explored the relationships between the different types of synaesthesia recorded by almost 20,000 people who completed the Synesthesia Battery (Eagleman, 2009) and found five distinct clusters of synaesthesia forms: coloured sequences (e.g. number-colours), coloured music (e.g. pitches-colours), coloured sensations (e.g. pain-colours), spatial sequences (e.g. calendar-forms), and non-visual sequelae (e.g. lexical-gustatory synaesthesia). In other words, Novich and colleagues (2011) observed that people who experienced more than one type of synaesthesia tended to experience other types that belonged in the same cluster (e.g. a person with letters-colours is more likely to additionally have numbers-colours than music-colours synaesthesia).

Since Novich et al.'s (2011) study was focused on -colour synaesthesias, we divided non-visual experiences or sequelae into specific categories based on the previous cited literature: -touch, -pain, -taste, -smell, -sound, and mirrors. We also renamed 'coloured music' to 'coloured sounds' to include other similar experiences within the cluster such as voices-colours. Lastly, we added three categories that we considered that were missing: visual patterns (i.e. -visual synaesthesias not related to -colour and characterised by the visualisation of forms such as lines, circles, or waves – with or without moment – in response to different triggers such as music or pain), ticker-tape synaesthesia (i.e. visualisation of spoken words or thoughts as 'subtitles' or like in a 'scrolling prompter'), and personifications (i.e. attribution of genders and/or personalities to concepts such as the letters, numbers, days of the week, etc.).

The final structure of the ESSA includes thus the following thirteen clusters or categories: coloured sequences, coloured sounds, coloured sensations, visual patterns, ticker-tape [synaesthesia], spatial sequences, -touch sequelae, -pain sequelae, -taste sequelae, -smell sequelae, -sound sequelae, mirror [synaesthesias], and personifications. Each category starts with a general statement that covers different triggers for that group of experience. For example, for the coloured sequences category: "I perceive different colours when I see, hear, or think about concepts such as letters, words, names, shapes, symbols, numbers, weekdays, months, hours, years, or measurement scales". In addition, an example or two are given for clarification purposes: e.g. "When I hear the word 'Tuesday', I perceive a canary yellow colour" or "When I think about the letter A, I perceive the colour red". If the responder states 'Not at all', the person is prompted to jump to the following category statement. If the responder gives any other answer (i.e. reports having this experience), then he or she is individually asked about all the possible different triggers within the category with specific statements that follow the same formulation: e.g. "I perceive different colours when I see, hear, or think about the numbers". The ESSA can be administered in this described detailed approach, covering all the category and applicable specific statements (Extended modality), or just going through the category statements (Brief modality).

In total, the ESSA has 13 category statements and 108 specific statements (coloured sequences 14, coloured sounds 5, coloured sensations 9, visual patterns 12, ticker-tape 1, synaesthesia 1, spatial sequences 8, -touch sequelae 7, -pain sequelae 7, -taste sequelae 12, -smell sequelae 12, -sound sequelae 11, mirror synaesthesias 4, and personifications 8). Ticker-tape synaesthesia is a specific type of synaesthesia which we considered that did not belong to any other category and thus the category statement acts as the specific one as well. The [category and specific] statements for the -touch and -pain sequelae experiences are shared, as we believed that they refer to similar experiences but of different degrees. Each statement asks about both sensations (e.g. “I experience physical touch or pain sensations in response to sounds/noises”) and if the responders report experiencing it, then they are further required to specify if it refers to -touch sensations, -pain, or both. The different experiences are later scored separately. Lastly, the screening tool includes a final open question to describe any additional experiences.

As mentioned, the ESSA can be administered as a questionnaire or as an interview, but here we focus on the validation of its usage as a questionnaire as we considered two main aspects with respect to its applicability to synaesthetic research: (a) potential usefulness (i.e. self-report aspect and the possibility of telematic administration), and (b) validation time (interview validation requires the involvement of independent raters and a series of additional protocols and steps).

4.2.2 Participants (Training and Validation samples)

The development and validation of the Edinburgh Synaesthesia Screening Assessment (ESSA) was conducted in parallel with the behavioural and personality studies. Participants of Studies 1 and 2 (total N = 149) were screened with the pilot version of the tool, which was used in the development of the measurement (see previous section). On the other hand, part of Study 3 participants completed the revised version of the ESSA and were included in the 'training' phase of the current study (Training sample). Lastly, an additional group of individuals were sampled for the 'testing' or 'validating' phase (Validation sample). All participants completed synaesthetic consistency tests for -colour and sequence-space synaesthesias. The data collection process of Study 5 was jointly conducted with Study 4, the Training and Validation samples respectively corresponding to Sample A and B (see section 3.2.2 in the previous Chapter for all the specifications).

The Training sample was composed of 31 synaesthetes (i.e. participants failing both consistency tests and not reporting any synaesthetic experiences), 86 synaesthetes (i.e. participants passing either or both consistency tests; in particular, 41 colour-synaesthetes, 18 sequence-synaesthetes, and 27 both-synaesthetes), and 39 other-synaesthetes (i.e. participants failing consistency tests for -colour and sequence synaesthetic experiences but reporting to have other types of synaesthesias that could not be assessed). In the Validation sample there were 134 non-synaesthetes, 62 synaesthetes (52 colour-synaesthetes and 10 sequence-synaesthetes; there were no synaesthetes experiencing both types), and 79 other-synaesthetes (in this subsample, i.e. participants failing both consistency tests, reporting to have other types of synaesthesias that could not be assessed, and passing the score thresholds for the group defined by the Training sample analyses; see section 4.3.2.1). Table 10 shows the main descriptive statistics for Study 5 samples and Table 7 in section 3.2.2 in the previous Chapter a summary of the average consistency scores obtained per group and sample[33].

[33] To assess the adequacy of the Training and Validation samples sizes, we consulted the work of Bujang and Adnan (2016), who calculated the requirements for minimum sample size for sensitivity and specificity analyses depending on different prevalence estimates. The Training sample of the present study had a [synaesthetic] prevalence of 73.5% (synaesthetes N = 86 and non-synaesthetes N = 31) and the Validation sample of 31.6% (synaesthetes N = 62, non-synaesthetes N = 134). According to the authors, to achieve a minimum power of .80 (actual power = .81) for detecting a change in the percentage value of sensitivity of a screening test from .50 to .70 based on a target significance level of .05 (actual p = .044), a minimum sample size of 70 subjects (including 49 having the condition) is needed when the prevalence of a disease or condition is estimated to be 70% and a minimum sample size of 163 subjects (including 49 having the condition) when the estimated prevalence is 30%. Thus, our proposed Training and Validation samples should be adequate for the aims of this study.

Table 10.
Descriptive, chi-square (χ^2), and t-*statistics of Study 5 groups' demographics.*

Demographics	Synaesthetes	Other-synaesthetes	Non-synaesthetes	Statistics
Training sample				
N (male)	73 (13)	33 (6)	20 (11)	$\chi^2(2) = 6.58$, $p = .037$
Age (SD)	21.6 (2.42)	22.1 (3.06)	22.5 (3.06)	$F(2, 153) = 1.71$, $p = .19$
Handedness (left, ambidextrous)	75 (6, 5)	34 (5, 0)	29 (2, 0)	$\chi^2(2) = 5.45$, $p = .24$
Nº of (native) languages* (SD)	1.27 (.52)	1.28 (.65)	1.13 (.34)	$F(2, 153) = .93$, $p = .40$
Level of education** (SD)	2.46 (.56)	2.54 (.60)	2.42 (.56)	$F(2, 153) = .55$, $p = .58$
Validation sample				
N (male)	50 (12)	59 (20)	111 (23)	$\chi^2(2) = 2.09$, $p = .35$
Age (SD)	21 (3.93)	21.4 (4.69)	21 (5.25)	$F(2, 270) = .20$, $p = .82$
Handedness (left, ambidextrous)	54 (6, 2)	72 (6, 1)	117 (17, 0)	$\chi^2(2) = 5.45$, $p = .24$
Nº of (native) languages* (SD)	1.05 (.22)	1.22 (.47)	1.12 (.35)	$F(2, 270) = 3.75$, $p = .025$
Level of education** (SD)	2.34 (.63)	2.19 (.56)	2.17 (.51)	$F(2, 270) = 2.05$, $p = .13$

Note: *N* = Sample size, SD = Standard Deviation.
* Nº of (native) languages: 1 = Monolingual, 2 = Bilingual, 3 = Polylingual.
** Level of education: 1 = High School, 2 = Undergraduate, 3 = Master, 4 = PhD, 5 = Postdoc.

As specified in the previous study, we used the loose synaesthetic consistency criteria (see section 3.2.2) to classify participants in both Studies 4 and 5 with the aim to capture as much as possible the complex reality and variability of the synaesthetic population, as we considered these key factors on the two topics covered (group and individual differences in personality traits and development of a screening tool). Concerning Study 5, omitting these individuals would imply a loss in population representation and, consequently, in screening power. Other-synaesthetes were also considered in the analyses for the same reasons, but given their more uncertain nature (since we could not assess the veracity of their experiences), the approach was different. In this case, we first conducted the training and validation analyses without them and then we run additional analyses including them as synaesthetes, evaluating their performance and comparing both results.

4.2.3 Data analyses

To validate the ESSA, we first scored the responses of all participants in the questionnaire transforming the 5-point Likert rating options as follows: 0-'Not at all', 1-'A little bit', 2-'Moderately', 3-'Quite a lot', and 4-'Completely'. If a person gave a 'Not at all' answer to a category statement, all the individual statements received a 0 as well. Although the -touch and -pain sequelae categories were asked together, individual scorings for each sensation were coded according to the responder specifications (see section 4.2.1 and the detailed instructions specified in the copy of the measurement in Appendix D).[34]

4.2.3.1 Training analyses.

We assessed the screening or predictive power of ESSA on the Training sample evaluating the optimal thresholds to distinguish between synaesthetes and non-synaesthetes. With that aim, several scoring systems were calculated and analysed for the Brief (administration of category or general statements only) and Extended (including the specific or detailed statements; see section 4.2.1) modalities. Given that from a theoretical standpoint there were no priors as to what constitutes a good method to quantify the synaesthetic experience, we devised four scoring strategies: sum of ratings, highest rating, average rating, and number of experiences (see Table 11 for details).

Table 11.
Specifications regarding the different Edinburgh Synaesthesia Screening Assessment scorings calculated and analysed for the Brief and Extended modalities in Study 5.

Scoring name	Abbreviation	Scoring calculation
Brief modality		
Sum	BS	Sum of the direct ratings for each category statement (range 0-52)
Highest	BH	Highest rating out of the direct ratings for each category statement (range 0-4)
Average	BA	Average of the direct ratings for all the category statements (range 0-4)
Number (of categories)	BN	Total number of category statements with direct ratings > 0 (range 0-13)
Extended modality		
Sum	ES	Sum of the direct ratings for each specific statement (range 0-432)
Highest	EH	Highest rating out of the direct ratings for each specific statement (range 0-4)
Average	EA	Average of the direct ratings for all the specific statements (range 0-4)
Number (of types)	EN	Total number of specific statements with ratings > 0 (range 0-108)

[34] Control analyses were conducted to disregard any possible influences of gender in the Training sample and of number of languages in the Validation sample (see Table 10). Firstly, we compared the ESSA scorings and direct statements' means of each group of the Training sample (non-synaesthetes, synaesthetes, and other-synaesthetes) between females and males with non-parametric Mann-Whitney Us (as most variables violated the assumption of normality, as assessed by Shapiro-Wilk). All analyses showed no mean differences in the scorings and statements' by gender for any of the groups (non-synaesthetes: all $U(29) > 81$, all $p > .12$; synaesthetes: all $U(84) > 289$, all $p > .011$; other-synaesthetes: all $U(37) > 73.5$, all $p > .31$; Bonferroni-adjusted alpha = .004). Regarding the differences in number of languages observed for the different groups of the validation sample, we submitted the different dependent variables of interest to linear regressions with number of languages as the predictor, separately for each group. The analyses revealed that number of languages did not predict mean scores for the scorings and statements assessed for any of the groups (non-synaesthetes: all $F(1, 132) < 3.70$, all $p > .056$, all $R^2 < .027$; synaesthetes: all $F(1, 60) < 5.68$, all $p > .020$, all $R^2 < .087$; other-synaesthetes: all $F(1, 77) < 2.87$, all $p > .094$, all $R^2 < .036$; Bonferroni-adjusted alpha = .004).

We then applied predictive ROC (Receiver Operating Characteristic) curve analyses to the binary classification of synaesthetes (without other-synaesthetes) and non-synaesthete controls according to consistency measures (loose criteria; see section 2.2.2.1.1). These analyses allowed us to determine the cut-off values maximising sensitivity (SE) and specificity (SP) for each scoring. A perfect screening questionnaire has the potential to completely discriminate subjects with and without the condition. In other words, in a perfect questionnaire, all the subjects with the condition have questionnaire values above the cut-off (true positives) and none below (false negatives), and all the subjects without the condition have questionnaire values below the cut-off (true negatives) and none above (false positives). Therefore, SE is defined as the percentage of true positives in a total group of subjects with the condition (sum of true positives and false negatives), and SP as the percentage of true negatives in a total group of subjects without the condition (sum of true negatives and false positives). Besides SE and SP, we used additional measures to compare and evaluate the different maximised output models for each scoring. Given that the class distribution of our sample was skewed (see Table 10 in section 4.2.2), we disregarded commonly used measures dependent on condition prevalence (e.g. diagnostic accuracy or positive and negative predictive values; e.g. Šimundič, 2009). Instead, we evaluated the following diagnostic measures: area under the curve (AUC), positive and negative likelihood ratios (PLR and NLR), diagnostic odds ratio (DOR), and Youden's J statistic (J).

The ROC curve emerges from plotting the different SE and SP values associated with each possible threshold or cut-off point. The closer the curve is located to the upper-left hand corner and, therefore, the larger the area under the curve (AUC), the better the classifier is at discriminating between the condition and the non-condition. A perfect classifier would have an AUC of 1.0 and a non-discriminating classifier would have an area of 0.5 (e.g. Fawcett, 2006; Haijan-Tilaki, 2013; Šimundič, 2009). The likelihood ratios tell us how many times more likely a particular classifier result is in subjects with the condition than those without it: the positive likelihood ratio (PLR) refers to positive results for people with the condition and the negative likelihood ratio (NLR) to negative results for people without the condition. Good diagnostic tests have PLR > 10 and NLR < 0.1 (e.g. Chou, Dana, & Bougatsos, 2009; Raslich, Markert, & Stutes, 2007; Šimundič, 2009). Similarly, the diagnostic odds ratio (DOR) is defined as the ratio of the odds of a classifier being positive if the subject has a condition relative to the odds of the classifier being positive if the subject does not have the condition. DOR ranges from 0 to infinity; the higher the SE and SP of a classifier (i.e. low rates of false positives and negatives), the higher the DOR (e.g. Glas, Lijmer, Prins, Bonsel, & Bossuyt, 2003; Raslich et al., 2007; Šimundič, 2009). Finally, Youden's J Statistic or Youden's Index (J) is a global diagnostic measure which assesses the overall discriminative power of a classifier. It is calculated by deducing 1 from the sum of the classifier's SE and SP; values ranging from 0 (non-diagnostic accuracy) to 1 (perfect diagnostic accuracy) (Youden, 1950).

After choosing the scoring models with the best overall performance for the Brief and the Extended modalities according to the evaluating measures detailed, we performed additional analyses to further assess the optimal scoring models. First, we were interested in evaluating whether other-synaesthetes subjects could also be discriminated from non-synaesthete controls with the same cut-offs provided by the chosen models. Therefore, we performed ROC curve analyses on the scoring models with a new sample that included other-synaesthetes in the synaesthete group. Second, we wanted to assess how these optimal general scorings compared to the predictive potential of the direct scores of those category or specific statements which are directly associated with synaesthetic consistency tests (e.g. Q2 for letters-colours). To do this, we submitted the direct scores of these statements to ROC curve analyses with binary classifications of presence/absence of the specific synaesthesia types according to each particular consistency test scores. We could only perform these analyses for those statements related to consistency tests for which we had complete consistency data (i.e. consistency data for all participants; Table 12). It is worth noticing that in these analyses all the sample was included without the [synaesthetes vs. other-synaesthetes] variable, since the distinction was simply between having or not having a specific synaesthetic experience.

Table 12.
Specifications regarding the different Edinburgh Synaesthesia Screening Assessment direct statements analysed in Study 5.

ESSA Statement	Type of statement	Synaesthetic consistency measure
Q1 Coloured sequences	Category	Synesthesia Battery / Multisense Consistency Test -colour score*
Q2 Letters-colours	Specific	Synesthesia Battery / Multisense Consistency Test Letters-colours score
Q46 Spatial sequences	Category	Sussex´s Sequence-Spatial Synaesthesia Diagnostic Test consistency score
Q48 Number-forms	Specific	Sussex´s Sequence-Spatial Synaesthesia Diagnostic Test consistency score
Q49 Weekday-forms	Specific	Sussex´s Sequence-Spatial Synaesthesia Diagnostic Test consistency score
Q50 Month-forms	Specific	Sussex´s Sequence-Spatial Synaesthesia Diagnostic Test consistency score

* All participants completed the letters-colours consistency test. However, the Training sample participants were also tested with other available -colour consistency tests (i.e. numbers, weekdays, months) if they reported having these experiences; the minimum score in any of these tests was used as reference to classify them and to perform the analyses regarding Q1.

4.2.3.2 Validation analyses.

Data validation analyses were conducted to ratify that the optimal models would perform similarly under different testing and sampling conditions. We adopted two methodological approaches to that aim: cross-validation analyses (on the Training sample) and external validation in a new sample (Validation sample).

Cross-validation is a method used to assess the predictive ability of a model and relies on the principle of splitting the data into two or more sub-groups which are used to train and test/validate. In particular, we run 10-fold cross-validation analyses on the Brief and Extended optimal models (with and without other-synaesthetes) and on each of the individual statements' models derived from the Training sample ROC analyses. The technique involves randomly partitioning the data sample into complementary subsets and performing the analysis on one subset (called the training set) and validating the analysis on the other subset (validation or testing set). To reduce variability, this analysis is repeated a number of times or folds using different partitions, and the validation results are averaged over the folds to give an estimate of the model's predictive performance (e.g. James, Witten, Hastie, & Tibshirani, 2013; Kuhn & Johnson, 2013). The evaluation metrics for this procedure were the above described SE, SP, and AUC, in addition to Cohen's Kappa (K), a classification accuracy measure normalised at the baseline of random chance on a given dataset that is especially useful for imbalanced samples (e.g. McHugh, 2012; Viera & Garrett, 2005).

Following next, we conducted external validation analyses, which consists on testing the model on a different sample of subjects (Validation sample here) and is considered a more robust approach (e.g. Salciccioli, Crutain, Komorowski, & Marshall, 2016). To do this, synaesthetes and non-synaesthetes (as defined by consistency measures; see section 2.2.2.1.1) were classified according to the Training sample optimal thresholds. That is, whether participants' scores on the ESSA failed or passed the predicted synaesthetic cut-offs defined in the ROC curve analyses. The resulting contingency tables (i.e. actual synaesthetic condition: yes/no; predicted synaesthetic condition: yes/no) were then submitted to the same evaluating measurements and indexes described in the previous section (i.e. sensitivity, specificity, PLR, NLR, DOR, and J). These analyses were performed with and without other-synaesthetes (included in the synaesthete group). To determine the other-synaesthetes of this sample, we selected those participants who did not pass both consistency tests but had (both Brief and Extended) scores above the mean scores obtained by other-synaesthetes in the Training sample (see Table 16 in section 4.3.2.1).

All the training and validation analyses were performed in R 3.5.1 (R Core Team, 2018) with the following packages: *stats* (R Core Team, 2018), *epiR* (Stevenson, 2019), *pROC* (Robin et al., 2011), and *cutpointr* (Thiele, 2018). The packages *plotROC* (Sachs, 2017), *ggplot2* (Wickham, 2016), *lattice* (Sarkar, 2008), and *caret* (Kuhn, 2019) were also used.

4.3 Results

4.3.1 ESSA responses overview

Table 13 details the mean scores for each defined ESSA scoring per group and sample and Figure 18 depicts the percentage of participants per group in the Training and Validation samples reporting to experience synaesthesia types in each category (i.e. coloured sequences, coloured sounds, etc.).

Table 13.
Mean scores by sample and group for the different Edinburgh Synaesthesia Screening Assessment scorings of Study 5.

	Training sample			**Validation sample**		
Scoring name	S (*N* = 86)	OS (*N* = 39)	NS (*N* = 31)	S (*N* = 62)	OS (*N* = 79)	NS (*N* = 134)
Brief modality						
Brief Sum (BS)	15.3 (9.94)	10.8 (8.38)	4.90 (5.60)	5.89 (5.73)	12.2 (6.43)	2.57 (2.45)
Brief Highest (BH)	2.99 (1)	2.54 (1.07)	1.58 (.96)	1.92 (1.26)	3.01 (.79)	1.05 (.74)
Brief Average (BA)	1.87 (.64)	1.61 (.59)	1.22 (.53)	1.29 (.75)	1.90 (.66)	.88 (.55)
Brief Number (BN)	7.42 (3.36)	6.05 (3.01)	3.35 (2.87)	3.47 (2.73)	6.52 (2.79)	2.13 (1.85)
Extended modality						
Extended Sum (ES)	91.5 (70.3)	57.2 (62.6)	21.8 (34.5)	22.2 (27.1)	55 (41.3)	6.06 (6.76)
Extended Highest (EH)	3.41 (.85)	2.92 (1.01)	1.87 (1.09)	2.06 (1.41)	3.37 (.62)	1.22 (.97)
Extended Average (EA)	2.29 (.56)	2.10 (.90)	1.57 (1.01)	1.28 (.73)	1.88 (.52)	.91 (.65)
Extended Number (EN)	37.9 (25.8)	25.6 (23.8)	10.8 (14.8)	13 (14.5)	29 (18.2)	4.82 (5.38)

Note: N = Sample size; Standard Deviation in parentheses; S = Synaesthetes; OS = Other-synaesthetes; NS = Non-synaesthetes.

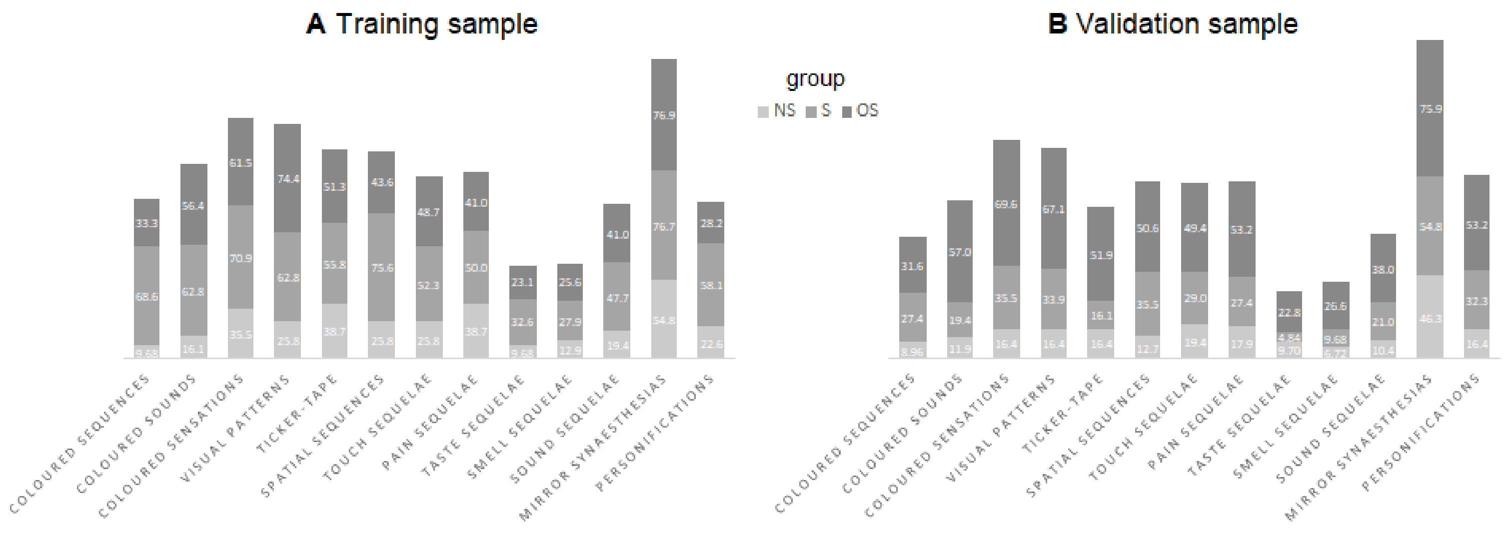


Figure 18. Percentile distribution of Study 5 participants reporting experiences in each specific synaesthetic category in the Edinburgh Synaesthesia Screening Assessment (ESSA), for each sample and group (non-synaesthetes – NS; synaesthetes – S; other-synaesthetes – OS).

4.3.2 Training analyses

We first applied predictive ROC (Receiver Operating Characteristic) curve analyses to the different ESSA scorings calculated and the binary classification of synaesthetes (without other-synaesthetes) and non-synaesthetes of the Training sample according to synaesthetic consistency measures. Synaesthetes showed significantly higher scores than non-synaesthetes irrespective of the ESSA scoring used (Table 14).

Table 14.
Mean scores by group for the different Edinburgh Synaesthesia Screening Assessment scorings of Study 5 (Training sample without other-synaesthetes).

Scoring name	S (N = 86)	NS (N = 31)	Statistics
Brief modality			
Brief Sum (BS)	15.3 (9.94)	4.90 (5.60)	z = 4.13, p < .001, d = 1.29
Brief Highest (BH)	2.99 (1)	1.58 (.96)	z = 4.91, p < .001, d = 1.44
Brief Average (BA)	1.87 (.64)	1.22 (.53)	z = 4.02, p < .001, d = 1.11
Brief Number (BN)	7.42 (3.36)	3.35 (2.87)	z = 4.62, p < .001, d = 1.55
Extended modality			
Extended Sum (ES)	91.5 (70.2)	21.8 (34.5)	z = 4.01, p < .001, d = 1.26
Extended Highest (EH)	3.41 (.85)	1.71 (1.07)	z = 5.45, p < .001, d = 1.76
Extended Average (EA)	2.29 (.56)	1.67 (1.01)	z = 3.80, p < .001, d = .76
Extended Number (EN)	37.9 (25.8)	10.8 (14.8)	z = 4.18, p < .001, d = 1.29

Note: N = Sample size; Standard Deviation in parentheses; S = Synaesthetes; NS = Non-synaesthetes.

Regarding the comparison of the different scoring models, the Brief Sum (BS) and Extended Highest (EH) offered the best overall performances as determined by the different evaluation metrics considered (Table 15 and Fig. 19). The BS or sum of the direct ratings for the different category statements suggested a cut-off of ≥ 6 points (score range: 0-52) (Fig. 20A). On the other hand, the optimal threshold for the EH or the highest rating out the direct ratings for each specific statement was ≥ 3, which corresponds to the questionnaire rating of 'Quite a lot' in the 5-point Likert scale ranging from 0-'Not at all' to 4-'Completely' (Fig. 20B).

Table 15.
Evaluation metrics for the different Edinburgh Synaesthesia Screening Assessment scorings of Study 5 (Training sample without other-synaesthetes).

Scoring name	Cut-off	SE	SP	AUC	PLR	NLR	DOR	J
Brief modality								
Brief Sum (BS)	≥ 6	83.7	74.2	85.2	3.24	.22	14.8	.58
Brief Highest (BH)	≥ 3	69.8	87.1	83.4	5.41	.35	15.6	.57
Brief Average (BA)	≥ 1.38	79.1	71	78.6	2.72	.29	9.23	.50
Brief Number (BN)	≥ 7	64	87.1	82	4.96	.41	12	.51
Extended modality								
Extended Sum (ES)	≥ 25	82.6	80.6	86.2	4.27	.22	19.7	.63
Extended Highest (EH)	≥ 3	87.2	77.4	87.7	3.86	.17	23.4	.65
Extended Average (EA)	≥ 1.85	83.7	58.1	76	2	.28	7.12	.42
Extended Number (EN)	≥ 13	81.4	77.4	84.4	3.60	.24	15	.59

Note: SE = Sensitivity; SP = Specificity; AUC = Area Under the Curve; PLR = Positive Likelihood Ratio; NPR = Negative Likelihood Ratio; DOR = Diagnostic Odds Ratio; J = Youden's J Statistic.

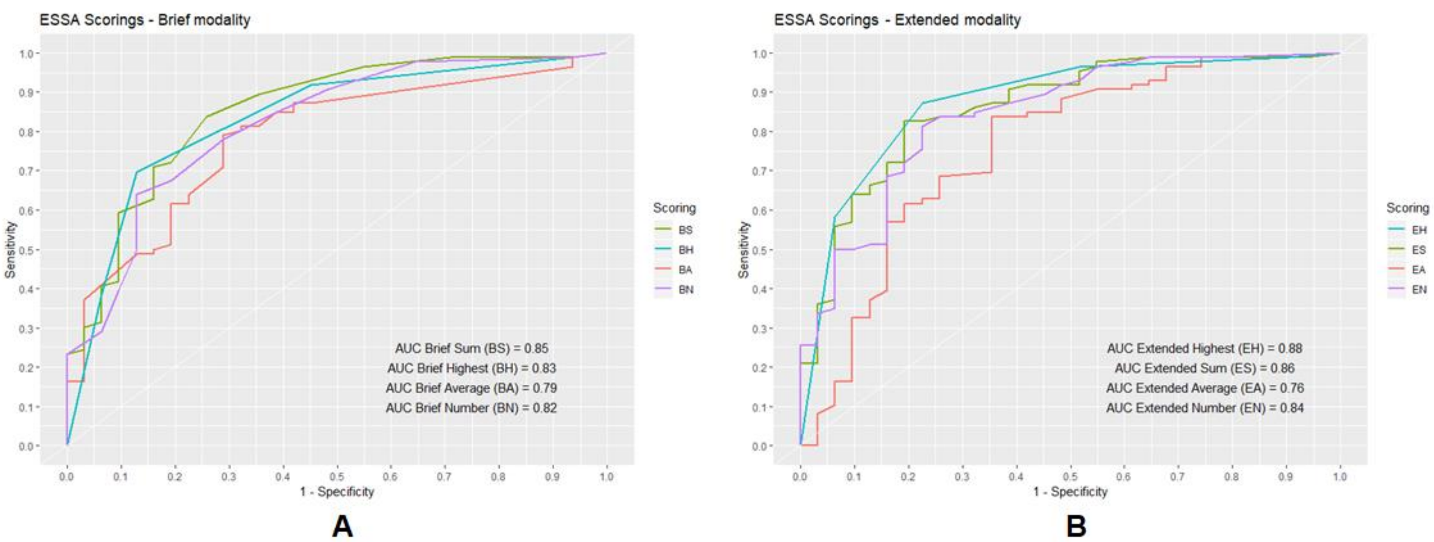
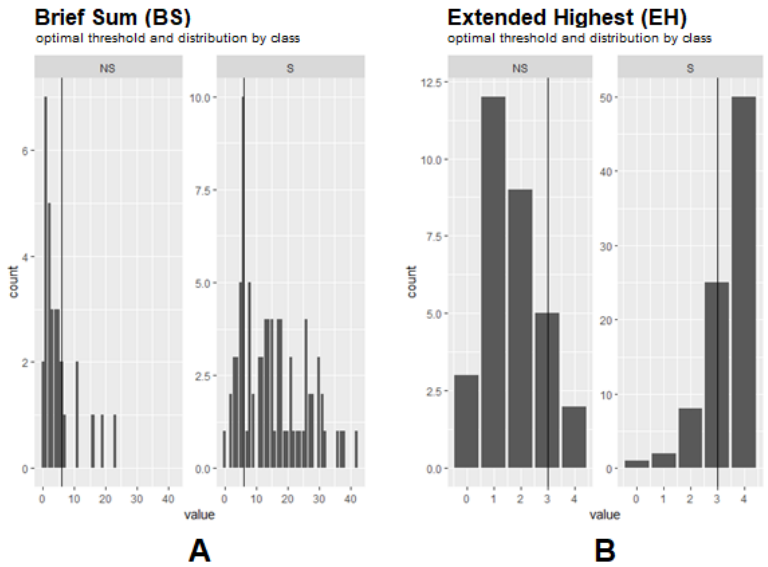


Figure 19. ROC (Receiver Operating Characteristic) curves and their associated AUC (Area Under the Curve) values for the different Brief (Figure **A**) and Extended (Figure **B**) scorings of the Edinburg Synaesthesia Screening Assessment (ESSA) – Study 5 Training sample without other-synaesthetes.



	S	NS
BS ≥ 6	72	8
BS < 6	14	23

	S	NS
EH ≥ 3	75	7
EH < 3	11	24

Figure 20. Optimal thresholds and distribution by synaesthetic class (non-synaesthetes – NS; synaesthetes – S) for the Brief Sum (BS; Figure **A**) and Extended Highest (EH; Figure **B**) scorings of the Edinburg Synaesthesia Screening Assessment (ESSA) – Study 5 Training sample without other-synaesthetes.

4.3.2.1 Other-synaesthetes analyses.

To evaluate whether other-synaesthetes could also be discriminated with the same cut-offs, we performed the same analyses including these participants to the synaesthete group and comparing them to non-synaesthete controls. Although synaesthetes scored higher than other-synaesthetes (see Table 13 in section 4.3.1) on all the ESSA scorings and these differences were significant (all $U(125) < 1{,}274$, all $p < .028$), clear and significant differences were also observed when comparing non-synaesthetes vs. the new synaesthete group which included other-synaesthetes (Table 16).

Table 16.
Mean scores by group for the different Edinburgh Synaesthesia Screening Assessment scorings of Study 5 (Training sample with other-synaesthetes included in the synaesthete group).

Scoring name	S (N = 125)	NS (N = 31)	Statistics
Brief modality			
B16ief Sum (BS)	13.9 (9.68)	4.90 (5.60)	$z = 4.03$, $p < .001$, $d = 1.14$
Brief Highest (BH)	2.85 (1.04)	1.58 (.96)	$z = 4.85$, $p < .001$, $d = 1.27$
Brief Average (BA)	1.79 (.63)	1.22 (.53)	$z = 3.87$, $p < .001$, $d = .98$
Brief Number (BN)	6.99 (3.31)	3.35 (2.87)	$z = 4.61$, $p < .001$, $d = 1.18$
Extended modality			
Extended Sum (ES)	80.8 (69.6)	21.8 (34.4)	$z = 3.73$, $p < .001$, $d = 1.07$
Extended Highest (EH)	3.25 (.94)	1.71 (1.07)	$z = 5.57$, $p < .001$, $d = 1.53$
Extended Average (EA)	2.23 (.69)	1.57 (1.01)	$z = 3.85$, $p < .001$, $d = .76$
Extended Number (EN)	34.2 (25.7)	10.8 (14.8)	$z = 3.97$, $p < .001$, $d = 1.12$

Note: N = Sample size; Standard Deviation in parentheses; S = Synaesthetes; NS = Non-synaesthetes.

The Brief Sum (BS) and Extended Highest (EH) models were also the best performing ones (Table 17 and Fig. 21) and the suggested cut-offs for these scorings were identical to those with the sample without other-synaesthetes (Fig. 22).

Table 17.
Evaluation metrics for the different Edinburgh Synaesthesia Screening Assessment scorings of Study 5 (Training sample with other-synaesthetes included in the synaesthete group).

Scoring name	Cut-off	SE	SP	AUC	PLR	NLR	DOR	J
Brief modality								
Brief Sum (BS)	≥ 6	77.6	74.2	82.7	3	.30	9.96	.52
Brief Highest (BH)	≥ 3	64	87.1	80.5	4.96	.41	12	.51
Brief Average (BA)	≥ 1.36	72.8	71	75.4	2.51	.38	6.54	.44
Brief Number (BN)	≥ 5	74.4	71	80	2.56	.36	7.10	.45
Extended modality								
Extended Sum (ES)	≥ 25	77.6	80.6	83.2	4	.28	14.4	.58
Extended Highest (EH)	≥ 3	80.8	77.4	84.6	3.58	.25	14.4	.58
Extended Average (EA)	≥ 1.85	76	64.6	72.2	2.14	.37	5.76	.41
Extended Number (EN)	≥ 13	76.8	77.4	81.4	3.40	.30	11.3	.54

Note: SE = Sensitivity; SP = Specificity; AUC = Area Under the Curve; PLR = Positive Likelihood Ratio; NPR = Negative Likelihood Ratio; DOR = Diagnostic Odds Ratio; J = Youden's J Statistic.

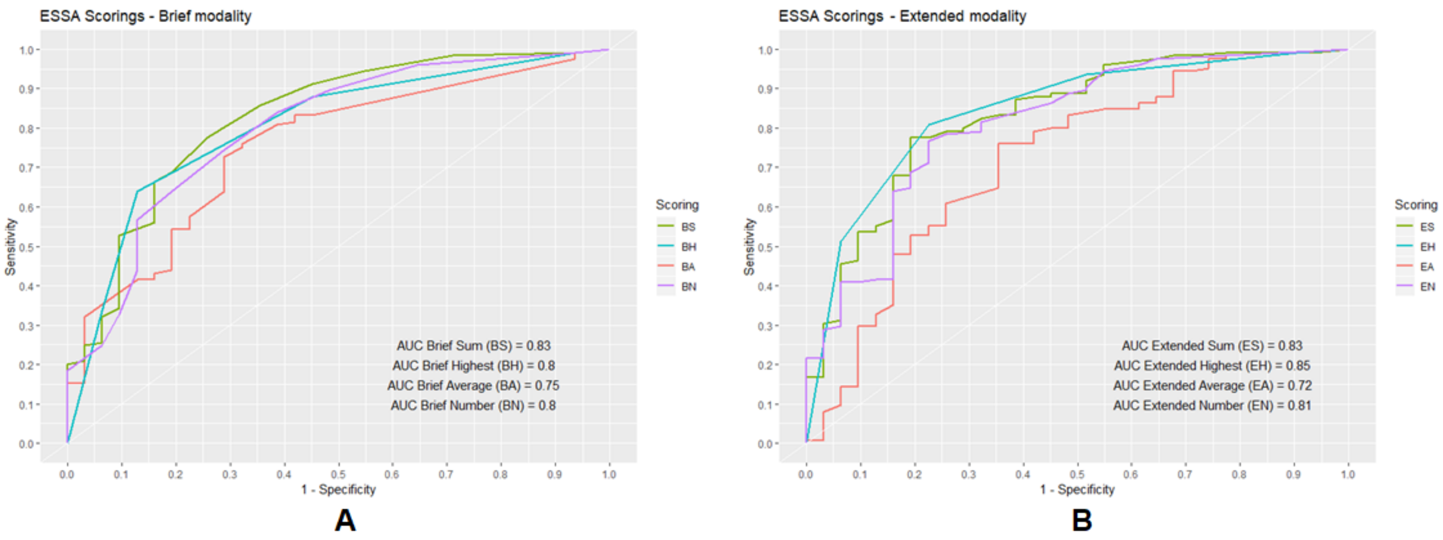
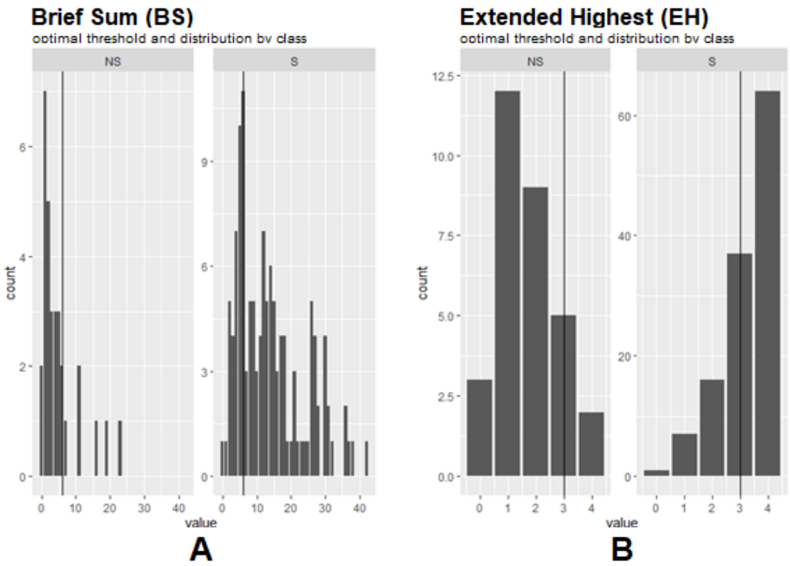


Figure 21. ROC (Receiver Operating Characteristic) curves and their associated AUC (Area Under the Curve) values for the different Brief (Figure **A**) and Extended (Figure **B**) scorings of the Edinburg Synaesthesia Screening Assessment (ESSA) – Study 5 Training sample with other-synaesthetes (included in the synaesthete group).



	S	NS
BS ≥ 6	97	8
BS < 6	28	23

	S	NS
EH ≥ 3	101	7
EH < 3	24	24

Figure 22. Optimal thresholds and distribution by synaesthetic class (non-synaesthetes – NS; synaesthetes – S) for the Brief Sum (BS; Figure **A**) and Extended Highest (EH; Figure **B**) scorings of the Edinburg Synaesthesia Screening Assessment (ESSA) – Study 5 Training sample with other-synaesthetes (included in the synaesthete group).

4.3.2.2 Category and specific statements analyses.

We also assessed the predictive power of the direct ratings of the category and specific statements which are directly associated with consistency tests (e.g. Q2 for letters-colours; see Table 12 in section 4.2.3.1). In these analyses, all participants were included and classified according to having or not each specific synaesthesia type (e.g. letters-colours:

yes/no). Thus, the general synaesthete vs. non-synaesthete or the synaesthetes with/without other-synaesthetes distinctions did not apply here (i.e. sequence synaesthetes were classified as a 'letters-colours: no' but this did not disqualify their synaesthetic condition in the e.g. 'number-forms: yes').

All the -colour and sequence category and specific statements analysed showed strong significant differences between people with and without the condition (Table 18). However, the performance of the direct ratings as synaesthetic classifiers were slightly below the general scoring models evaluated above (Table 19 and Fig. 23 and 24).

Table 18.
Mean scores by group for the different Edinburgh Synaesthesia Screening Assessment direct statements of Study 5 (Training sample).

Direct statement	W	WO	*N* W/WO	Statistics
Category				
Q1 Coloured sequences	1.49 (1.09)	.42 (.84)	68/88	$z = 5.46$, $p < .001$, $d = 1.10$
Q46 Spatial sequences	2.29 (1.26)	.85 (1.22)	45/111	$z = 5.33$, $p < .001$, $d = 1.16$
Specific				
Q2 Letters-colours	1 (1.20)	.42 (.84)	30/126	$z = 2.86$, $p = .004$, $d = .56$
Q48 Number-forms	2.18 (1.53)	.70 (1.18)	45/111	$z = 5.28$, $p < .001$, $d = 1.08$
Q49 Weekday-forms	2.58 (1.42)	.93 (1.34)	45/111	$z = 5.50$, $p < .001$, $d = 1.20$
Q50 Month-forms	2.58 (1.47)	.90 (1.37)	45/111	$z = 5.49$, $p < .001$, $d = 1.18$

Note: *N* = Sample size; Standard Deviation in parentheses; W = Participants with specific synaesthesia types; WO = Participants without specific synaesthesia types.

Table 19.
Evaluation metrics for the different Edinburgh Synaesthesia Screening Assessment direct statements of Study 5 (Training sample).

Direct statement	Cut-off	SE	SP	AUC	PLR	NLR	DOR	J
Category								
Q1 Coloured sequences	≥ 1	77.9	75	78	3.12	.29	10.6	.53
Q46 Spatial sequences	≥ 1	93.3	56.8	80.1	2.16	.12	18.4	.50
Specific								
Q2 Letters-colours	≥ 1	53.3	74.6	64.7	2.10	.63	3.36	.28
Q48 Number-forms	≥ 1	80	67.6	77.3	2.47	.30	8.33	.48
Q49 Weekday-forms	≥ 1	88.9	60.4	79.6	2.24	.18	12.2	.49
Q50 Month-forms	≥ 1	86.7	63.1	79.1	2.35	.21	11.1	.50

Note: SE = Sensitivity; SP = Specificity; AUC = Area Under the Curve; PLR = Positive Likelihood Ratio; NPR = Negative Likelihood Ratio; DOR = Diagnostic Odds Ratio; J = Youden's J Statistic.

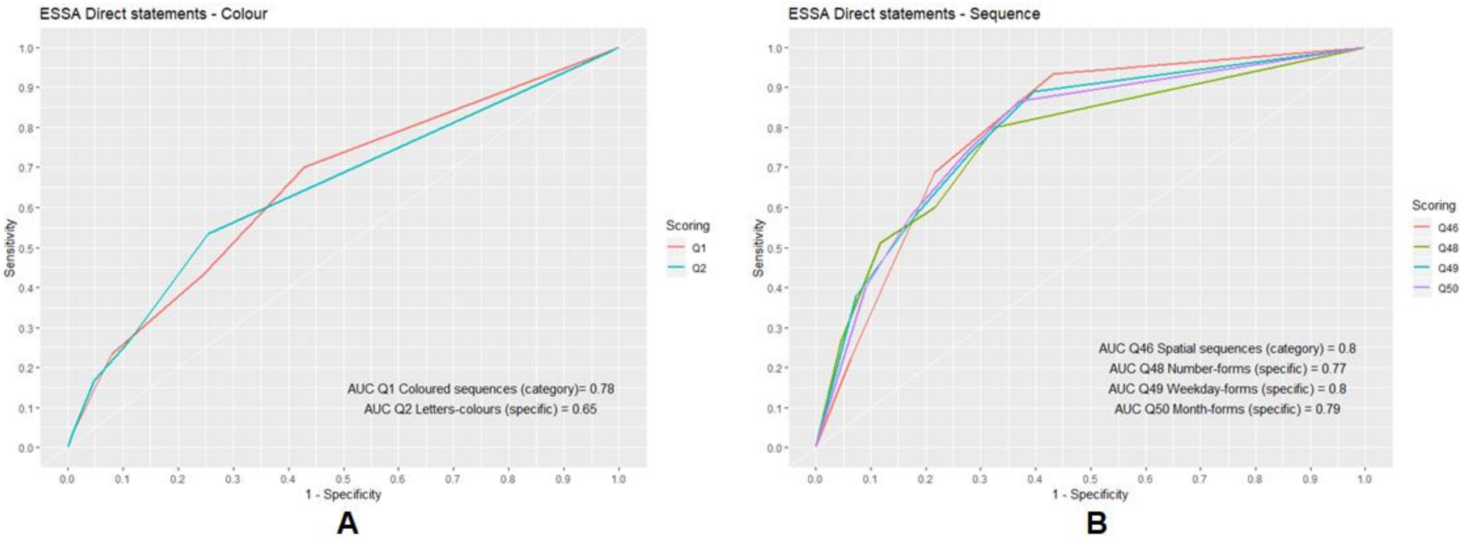


Figure 23. ROC (Receiver Operating Characteristic) curves and their associated AUC (Area Under the Curve) values for the different direct statements of the Edinburg Synaesthesia Screening Assessment (ESSA) related with -colour (Figure **A**) and sequence-spatial (Figure **B**) consistency scores – Study 5 Training sample.

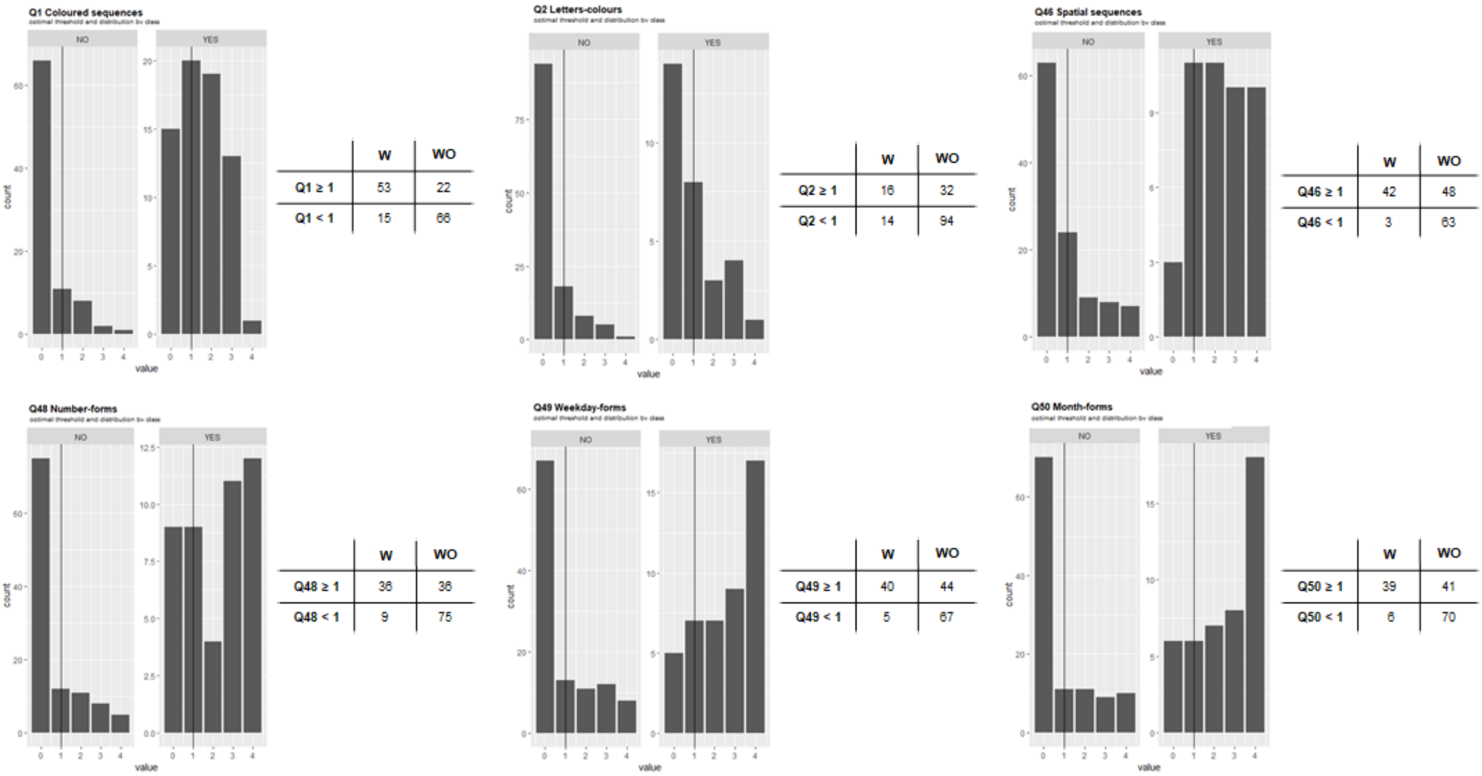


Figure 24. Optimal thresholds and distribution by synaesthetic class (participants with specific synaesthesia types – W; participants without specific synaesthesia types – WO) for different direct statements of the Edinburg Synaesthesia Screening Assessment (ESSA) related with -colour and sequence-spatial consistency scores – Study 5 Training sample.

4.3.3 Validation analyses

4.3.3.1 10-fold cross-validation analyses (Training sample).

In order to validate the Brief and Extended optimal models (with and without other-synaesthetes) and the direct ratings of the statements associated with particular synaesthetic consistency tests, we run 10-fold cross-validation analyses. The evaluation metrics for these analyses revealed low sensitivity and high specificity rates for the general scorings and vice-versa for the direct statements. Regarding Cohen's Kappa and according to Viera and Garrett (2005), the cross-validation results can be considered to have moderate to fair agreement for the general scorings and fair for the statements' direct ratings (except Q2 Letters-colours, with only slight agreement) (Table 20).

Table 20.
Evaluation metrics for the 10-fold cross-validation on the different Edinburgh Synaesthesia Screening Assessment scorings and direct statements of Study 5 (Training sample).

Scoring/Direct statement	SE	SP	AUC	K
Brief Sum (BS)	57.5	92.9	85.8	53.1
Brief Sum (BS) with other-synaesthetes	19.2	98.5	83.3	21.8
Extended Highest (EH)	46.7	93.2	88.7	40.7
Extended Highest (EH) with other-synaesthetes	46.7	93.2	88.7	40.4
Q1 Coloured sequences	87.5	49	78.4	37.8
Q46 Spatial sequences	86.6	44	79.2	31.7
Q2 Letters-colours	97.7	3.33	64.7	1.65
Q48 Number-forms	88.4	51	76.1	41.4
Q49 Weekday-forms	86.5	49.5	79.1	35.2
Q50 Month-forms	85.6	43	78.1	28.6

Note: SE = Sensitivity; SP = Specificity; AUC = Area Under the Curve; K = Cohen's Kappa.

4.3.3.2 External validation analyses (Validation sample).

We first analysed group differences in the Brief Sum and Extended Highest ESSA scorings and the particular ESSA statements directly associated with synaesthetic consistency scores. All results showed strong significant differences between synaesthetes and non-synaesthetes in the general scoring models. People with and without sequence synaesthesia obtained also significant score differences in the associated particular statements. However, this was not the case for the analyses concerning people with and without -colour synaesthesia and the specific ESSA statements related to this type (Table 21).

Table 21.
Mean scores by group for the different Edinburgh Synaesthesia Screening Assessment scorings and direct statements of Study 5 (Validation sample).

Scoring/Direct statement	S-W	NS-WO	*N* S-W/NS-WO	Statistics
Brief Sum (BS)	5.89 (5.73)	2.57 (2.45)	62/134	$z = 4.56$, $p < .001$, $d = .75$
Brief Sum (BS) with other-synaesthetes	9.43 (6.87)	2.57 (2.45)	141/134	$z = 7.72$, $p < .001$, $d = 1.33$
Extended Highest (EH)	2.06 (1.41)	1.22 (.97)	62/134	$z = 4.38$, $p < .001$, $d = .69$
Extended Highest (EH) with other-synaesthetes	2.79 (1.23)	1.22 (.97)	141/134	$z = 8.38$, $p < .001$, $d = 1.42$
Q1 Coloured sequences	.39 (.71)	.10 (.36)	52/223	$z = 1.164$, $p = .24$
Q46 Spatial sequences	.85 (1.40)	.16 (.46)	10/265	$z = 3.83$, $p < .001$, $d = .66$
Q2 Letters-colours	0 (0)	.01 (.02)	52/223	$z = .015$, $p = .99$
Q48 Number-forms	.60 (1.22)	.10 (.39)	10/265	$z = 3.84$, $p < .001$, $d = .55$
Q49 Weekday-forms	.84 (1.42)	.17 (.58)	10/265	$z = 3.72$, $p < .001$, $d = .62$
Q50 Month-forms	.87 (1.44)	.17 (.58)	10/265	$z = 3.69$, $p < .001$, $d = .64$

Note: *N* = Sample size; Standard Deviation in parentheses; S-W = Synaesthetes / People with specific synaesthesia types; NS-WO = Non-synaesthetes / People without specific synaesthesia types.

Then, after classifying the participants according to their synaesthetic status and the optimal thresholds defined in the ROC curve analyses of the training analyses, we evaluated the resulting contingency tables (i.e. actual synaesthetic condition: yes/no; predicted synaesthetic condition or scoring passing: yes/no; Figure 25). The metrics values were moderate for the general scorings (Brief Sum and Extended Highest) when they included other-synaesthetes and fair without them. The direct ratings for the sequence synaesthesias statements (Q46 Spatial sequences, Q48 Number-forms, Q49 Weekday-forms, Q50 Calendar-forms) also obtained moderate values. However, the -colour statements (Q1 Coloured sequences, Q2 Letters-colours) were poor to inexistent. See Table 22 for all the details.

ESSA Scorings

Brief Sum (BS)

	S	NS
BS ≥ 6	26	17
BS < 6	36	117

Brief Sum (BS)
with other-synaesthetes

	S	NS
BS ≥ 6	100	17
BS < 6	41	117

Extended Highest (BS)

	S	NS
EH ≥ 3	26	13
EH < 3	36	121

Extended Highest (BS)
with other-synaesthetes

	S	NS
EH ≥ 3	99	13
EH < 3	42	121

ESSA Direct statements

Q1 Coloured sequences

	W	WO
Q1 ≥ 1	12	42
Q1 < 1	40	181

Q2 Letters-colours

	W	WO
Q2 ≥ 1	0	1
Q2 < 1	52	222

Q46 Spatial sequences

	W	WO
Q46 ≥ 1	7	72
Q46 < 1	3	193

Q48 Number-forms

	W	WO
Q48 ≥ 1	7	48
Q48 < 1	3	217

Q49 Weekday-forms

	W	WO
Q49 ≥ 1	7	67
Q49 < 1	3	198

Q50 Month-forms

	W	WO
Q50 ≥ 1	7	65
Q50 < 1	3	200

Figure 25. Optimal thresholds and distribution by synaesthetic class (non-synaesthetes – NS; synaesthetes – S or participants with specific synaesthesia types – W; participants without specific synaesthesia types – WO) for the different scorings and the different direct statements related with -colour and sequence-spatial consistency scores of the Edinburg Synaesthesia Screening Assessment (ESSA) – Study 5 Validation sample.

Table 22.
Evaluation metrics for the external validation analyses on the different Edinburgh Synaesthesia Screening Assessment scorings and direct statements of Study 5 (Validation sample).

Scoring/Direct statement	SE	SP	PLR	NLR	DOR	J
Brief Sum (BS)	41.9	87.3	3.31	.67	4.97	.29
Brief Sum (BS) with other-synaesthetes	70.9	87.3	5.59	.33	16.8	.58
Extended Highest (EH)	41.9	90.3	4.32	.64	6.72	.32
Extended Highest (EH) with other-synaesthetes	70.2	90.3	7.24	.33	21.9	.61
Q1 Coloured sequences	23.1	70.2	1.23	.95	1.29	.04
Q46 Spatial sequences	70	72.8	2.58	.41	6.26	.43
Q2 Letters-colours	0	8.07	0	1	0	.0
Q48 Number-forms	70	81.9	3.86	.37	10.6	.52
Q49 Weekday-forms	70	74.6	2.77	.41	6.90	.45
Q50 Month-forms	70	75.5	2.85	.40	7.18	.45

Note: SE = Sensitivity; SP = Specificity; PLR = Positive Likelihood Ratio; NPR = Negative Likelihood Ratio; DOR = Diagnostic Odds Ratio; J = Youden's J Statistic.

4.4 Discussion

The study of synaesthesia rests on the comparison of synaesthetes and non-synaesthetes. This necessarily implies a clear definition of what does and does not constitutes synaesthesia. Synaesthetic tests of genuineness or consistency tests, which measure how consistent are people at reporting their specific synaesthetic associations over time or repeated trials, have proven to be robust and reliable in distinguishing between synaesthetes and non-synaesthetes. However, they are only available for a few synaesthesia types and this means that synaesthetic heterogeneity needs to be addressed with complementary approaches. Several studies have created self-report questionnaires or interviews for such purposes, but they are mostly suited to particular investigations needs; no systematic or standardised synaesthesia screening measurement exists to date. To address this gap in the synaesthetic research field, we developed the Edinburgh Synaesthesia Screening Assessment (ESSA), a self-report questionnaire/interview that aims to cover an exhaustive range of synaesthesia types and to assess both synaesthetes and non-synaesthetes.

The ESSA is organised in different categories of synaesthesia types that share the concurrents induced and/or common characteristics. For example, coloured sequences include those synaesthesias triggered by concepts such as letters, numbers, months, etc. which trigger -colour experiences; or the mirror synaesthesias category consists of experiences which involve feeling physical sensations like touch or pain in one's own body in response to seeing other people being touched or in pain. The ESSA has thirteen different categories: coloured sequences, coloured sounds, coloured sensations, visual patterns, ticker-tape synaesthesia, spatial sequences, -touch sequelae, -pain sequelae, -taste sequelae, -smell sequelae, -sound sequelae, mirror synaesthesias, and personifications. Each category starts with a descriptive statement which includes different inducers for that group of experiences (e.g. coloured sounds: "I perceive different colours when I listen to or hear music, sounds/noises, or voices"), together with examples (e.g. "When I listen to pop music, I see pink to red tonalities"). The person is then asked to rate how much does this experience applies to him/her on a 5-point Likert scale ranging from ‘Not at all’ to ‘Completely’. If the person reports having the experience (i.e. any response except ‘Not at all’), individual statements for each possible trigger within the category are presented (e.g. “I perceive different colours when I listen to or hear voices”). There is a total of thirteen category statements and 108 specific statements (i.e. possible types of synaesthesias). The ESSA offers two administration modalities: Brief – only the category statements, or Extended – category plus the individual specific statements (see section 4.2.1 and Appendix D for all the details).

In order to assess the predictive power of the ESSA (questionnaire version only), we evaluated the optimal thresholds to discriminate between synaesthetes and non-synaesthetes in a sample of 31 non-synaesthetes and 86 grapheme-colour and/or sequence-spatial synaesthetes verified via consistency tests (Training sample). To do that, participants’ ratings for each statement were transformed into scorings that followed different calculation approaches (sum of ratings, highest rating, average rating, and number of experiences), independently for the Brief and Extended administration modalities. Then, we applied predictive ROC (Receiver Operating Characteristic) curve analyses to the binary classification of synaesthetes and non-synaesthetes according to consistency measures (i.e. actual synaesthetic condition: yes/no; ESSA predicted synaesthetic condition: yes/no). The analyses showed that synaesthetes obtained significantly higher rates than non-synaesthetes in all the calculated scorings. However, the thresholds or cut-offs derived from the sum of ratings in the Brief modality (Brief Sum: synaesthetes ≥ 6 points – score range: 0-52) and in the highest rating in the Extended modality (Extended Highest: synaesthetes ≥ 3 points – score range: 0-4) had the best overall performances. That is, these scoring models offered the cut-off values with the greatest sensitivity (i.e. detection of synaesthetes) and specificity (i.e. detection of non-synaesthetes).

In particular, the Brief Sum scoring obtained a sensitivity of 83.7% and a specificity of 74.2%, whereas the Extended Highest scoring values were slightly higher: 87.2% and 77.4%, respectively. Although these rates are acceptable, they indicate some response biases. Examination of the overall pattern of responses in the ESSA showed that most participants reported experiencing some type of synaesthesia (i.e. they answered at least ‘A little bit’ to one or more statements). This includes as well non-synaesthetes, reflecting a certain degree of acquiescence or yes-saying response bias (i.e. tendency to answer ‘yes’ to all questions; e.g. Choi & Pak, 2005). On the other hand, 1.16% of synaesthetes said ‘Not at all’ to all the statements, suggesting a negligible no-saying response bias or proneness to answer ‘no’ to everything. Other assessed evaluation metrics seemed to reflect this as well. Positive and negative likelihood ratios (PLR and NLR), which inform about how many times a classifier is more likely to return the presence or absence of the condition for people with and without it, were below the optimal values (> 10 and < 0.1, respectively): Brief Sum – PLR = 3.24, NLR = .22; Extended Highest – PLR = 3.86, NLR = .17). Similarly, the scorings obtained only modest diagnostic odd ratios (DOR; i.e. ratio of the odds of a classifier correctly detecting the condition when is present relative to the odds of the classifier wrongly detecting it when is absent and it is the more higher the more sensitivity and specificity rates are close to 100%) and Youden’s J indexes (J; global diagnostic measure of the overall discriminative power of a classifier ranging from 0 or non-diagnostic accuracy to 1 or perfect diagnostic accuracy): Brief Sum – DOR = 14.8, J = .58; Extended Highest – DOR = 23.4, J = .65.

Following next, we evaluated the performance of the ESSA considering other-synaesthetes (i.e. 39 subjects not passing consistency tests but reporting to have other types of synaesthesias), who were removed from the main analyses to reduce noise. Interestingly, ROC curve analyses showed clear and significant differences between the scores of non-synaesthetes and synaesthetes (validated and other- grouped together here). Moreover, the Brief Sum and Extended Highest scorings were also the best performing ones and the analyses suggested the same synaesthetic cut-offs (≥ 6 and ≥ 3 points, respectively). On the other hand, we examined the predictive power of the direct ratings of the category and specific statements directly associated with consistency tests (i.e. grapheme-colour and sequence-spatial synaesthesias). Similarly to the previous analyses, all the statements of interest examined revealed significant higher scores for people with the specific synaesthesia types compared to people without, with suggested cut-offs of ≥ 1 point in this case. Although it should be noted that the evaluation metrics for these additional analyses were weaker than those for the main analyses (especially for the direct statements), these results seem to support the predictive power of ESSA to discriminate synaesthetes from non-synaesthetes.

Finally, we conducted several analyses to validate the observed results. First, we performed internal validation analyses on the Training sample by running a 10-fold cross-validation, which involves partitioning the sample into random training and validation subsets a number of times (10, in this case) and averaging the model's predictive performance results over these folds. We assessed the Brief Sum and Extended Highest optimal scorings (with and without other-synaesthetes) and the direct statements with this method. The analyses showed low sensitivity and high specificity rates for the general scorings and the opposite pattern for the direct statements. In addition, Cohen's Kappa inter-reliability metrics determined that there was fair to moderate agreement for the general scoring analyses and fair agreement for the direct statements ones. We also validated the optimal scoring models on an external sample composed of 134 non-synaesthetes, 62 grapheme-colour or sequence-space synaesthetes, and 79 other-synaesthetes (Validation sample; all participants completed the same synaesthetic consistency tests as the Training sample). The Brief Sum and Extended Highest scores of these participants were calculated and compared between groups. The analyses showed significant higher scores for synaesthetes compared to non-synaesthetes (with and without including other-synaesthetes). Significant group differences were also found between the scores of people who presented sequence synaesthesia and those who did not in relation to the direct statements associated with this type, but this was not the case for colour- synaesthetes and the respective -colour statements. Subjects were then classified according to their synaesthetic status (yes/no) and as passing or failing the cut-offs defined in the training phase for all the ESSA scorings and particular statements.

Sensitivity, specificity, PLR, NLR, DOR, and J values were examined. The results were moderate for the general scorings when they included other-synaesthetes and fair without them. The direct statements ratings were moderate for the sequence statements and poor to inexistent for the -colour ones.

The external validation results are thus somewhat inconsistent. One possibility is that this could be due to the fact that the no-saying bias was notably larger in the Validation sample, 14.5% of synaesthetes (mostly colour-synaesthetes) reporting 'Not at all' for all possible experiences[35]. This elevated percentage is especially important if we take into account that only 1.16% of synaesthetes had the same pattern of responses in the Training sample. Relatedly, while both synaesthetes and non-synaesthetes of the Validation sample obtained overall lower [Brief Sum and Extended Highest] scores compared to the Training sample, differences were especially evident for synaesthetes: Training sample (Brief Sum = 15.3, Extended Highest = 3.41), Validation sample (Brief Sum = 5.89, Extended Highest = 2.06). Differences between the two samples were already detected in the analyses regarding personality traits in Study 4 (both studies emerged from the same data collection process). As discussed there, although both samples were naïve to the purposes of the questionnaire and completed the ESSA on-line before doing any synaesthetic consistency tests, the Training participants came afterwards to the lab to complete additional tasks for other studies. We further hypothesised that this higher implication and commitment of the Training sample might inherently make them more motivated participants, explaining the higher scores of this sampled compared to the more conservative ones of the Validation sample.

.

[35] This elevated percentage of no-saying bias is mostly driven by the responses for the direct ESSA statement addressing letters-colours synaesthesia (Q2), with means of 0 or close to 0 from both people with and without the condition. We checked whether these results were caused by a technical issue (in the questionnaire administered or during the score calculation process), but we could not find anything backing this possibility. For that reason, we treated the results as valid. However, given their unlikeliness, it should be noted that we do not fully disregard the idea that we missed something.

As mentioned in Study 4 discussion as well, synaesthesia research is particularly prone to replicability issues due to difficulty to recruit big sample sizes and, most importantly, the lack of a well-established definition of what synaesthesia is and is not (see section 4.1. for further details on the issue), resulting in individual variability within subjects of different studies. One way of addressing this issue is to incorporate multiple measurements (e.g. consistency tests and behavioural tests such as the synaesthetic Stroop). But, as pointed out, there is the possibility that synaesthesia might be just very difficul to measure or quantify and, thus, that priority should be given to phenomenological approaches or more qualitative methodologies such as self-report questionnaires and interviews.

Taking this into account, one way in which the present study design could be improved, especially in relation to the recruited on-line sample (Sample B), would be to systematically ask for descriptions and examples for each type of self-reported synaesthetic experience. These could be then classified into different 'strength' of evidence according to previously set criteria. For instance, rejecting culturally-ingrained associations (e.g. "Cold temperatures are blue") as valid examples, or rating more synaesthetic-alike descriptions (as described in previous literature) more strongly (e.g. "I see the months of the year in an oval shape about half a meter to my right and I can zoom in and out to the current month") (see Rouw & Scholte, 2016 for an example on how to implement such a method). The description and examples ratings would be then incorporated to the overall criteria to decide which individuals are considered to experience synaesthesia and which not. In addition, implementing a test-retest for both the synaesthesia screening and the personality questionnaires with at least a week in-between each session, could provide a further measure that might be used to ponderate the classification of participants or to directly disregard or not particular subjects. In sum, the present study represents the first attempt to develop a systematic and standardised synaesthesia screening measurement. The Edinburgh Synaesthesia Screening Assessment (ESSA) was developed to capture as much as possible the broad synaesthetic spectrum, including an extensive number of synaesthesias (108 possible types), organised in thirteen categories. To make the test sensitive to different degrees of synaesthetic experience (including the no-experience for non-synaesthetes), the measurement uses a “How much does this [experience] applies to you?” formulation with a 5-point Likert response scale ranging from 0-‘Not at all’ to 4-‘Completely’. ROC curve analyses showed that a cut-off of ≥ 6 points (sum of ratings; 0-52) administering the questionnaire in its Brief version or ≥ 3 points (highest rating; range 0-4) in the Extended version significantly distinguishes (consistency verified) synaesthetes from non-synaesthetes, offering acceptable rates of sensitivity and specificity (±85.5 and ±75.8, respectively). However, the ESSA has shown some degree of yes-saying and no-saying response biases and modest validation results. Different combinations of cognitive, formatting, and psychometric approaches can be used to address these issues (e.g. Jobe, 2003). For example, the inclusion of additional, unrelated questions could be a strategy to control for the observed response biases (Chun & Hupé, 2016). Or reducing the number of items or statements of the questionnaire, analysing response frequencies and/or re-assessing literature fit, could help reduce questionnaire completion fatigue and frustration and, hopefully, yes-saying and no-saying biases. In addition, given the differences observed between the two samples assessed in this investigation, further validation studies with new samples would be desirable taking into account the above specified suggestions. Lastly, although the present study has only focused on the evaluation of the ESSA as a self-report questionnaire, it could also be interesting to consider the validation of its administration as an interview in future studies.

5. Chapter V: General Discussion

Synaesthesia is a perceptual condition in which stimulation of one sensory or cognitive pathway leads to automatic and involuntary experiences in a secondary sensory or cognitive pathway (e.g. seeing music or tasting words). Although synaesthetes experience additional percepts during their inducer-concurrent associations that are often unrelated or irrelevant to their daily activities, they appear to be relatively unaffected by this potentially distracting information. Chapter II investigated whether -visual synaesthetes (i.e. those experiencing at least one synaesthesia type involving visual concurrents or triggers such as grapheme-colour or sequence-space synaesthesias) were better than-non synaesthetes at filtering out other non-synaesthetic irrelevant visual stimuli in different conflict tasks. In Study 1, participants were asked to complete a visuo-tactile cross-modal congruency task focusing on tactile targets and ignoring visual distractors. In addition, they performed a visual, non-spatial congruency task (Flanker task). In Study 2 participants completed the same visuo-tactile task of Study 1, but in this case the instructions regarding the target and distractor sensory modalities were reversed and, thus, individuals had to attend to visual targets whilst ignoring tactile distractors. In addition, participants performed a unimodal visual version of this task which presented different visual targets and distractors. Synaesthetes (and particularly those who experienced sequence-space synaesthesias) were more efficient than controls at ignoring the irrelevant stimuli of the visuo-tactile task in which vision was task-irrelevant (i.e. task which presented visual distractors and tactile targets; Study 1); no other group differences emerged.

However, these results were not replicated in Study 3, which assessed a new sample of -visual synaesthetes with the two versions of these visuo-tactile task (i.e. vision task-relevant and vision task-irrelevant). In addition, we also evaluated whether the observed synaesthetic attentional advantages were consistent across different sensory modalities combinations. To that aim, participants performed audio-visual equivalents of the same visuo-tactile tasks (i.e. audio-visual stimuli in which vision was task-relevant in one task and task-irrelevant in the other). Furthermore, we investigated whether different types of -visual synaesthetes showed the same abilities or not by specifically comparing matched groups of non-synaesthetes, colour-synaesthetes (i.e. -visual synaesthetes who experienced synaesthesias involving -colour as the concurrent; e.g. grapheme-colour synaesthesia), and sequence-synaesthetes (i.e. -visual synaesthetes who experienced sequence-space synaesthesia forms; e.g. calendar-forms). Results revealed that sequence-synaesthetes were better than non-synaesthetes *and* colour-synaesthetes at filtering tactile irrelevant distractors presented with visual targets (visuo-tactile task, vision task-relevant version); we did not observe any other significant differences between groups. To try to understand the discrepancies between the different studies' results, we conducted a mini meta-analysis pooling the diverse effect sizes for the non-synaesthete vs. synaesthete (colour-, sequence-, or both, depending on the study) comparisons of the different studies of Chapter II. The analyses revealed a low overall true effect size and between-study heterogeneity, but they did not suggest the exclusion of any particular study (see section 2.5 for details). Therefore, the reason for the contrasting results observed might be a more intricate enterprise.

The findings of Study 3 importantly suggest that experiencing specific types of synaesthesias might play a relevant role in shaping the cognitive abilities of synaesthetes. This was further supported by a series of exploratory analyses that involved the different studies of Chapter II. Studies 1 and 2 samples were composed of different types of synaesthetes, including colour- and sequence-synaesthetes and people experiencing these two types of synaesthesia (both-synaesthetes). In Study 3, both-synaesthetes were excluded from the main analyses because the aim was precisely to assess differences between types, but we later asked the question of how these individuals having both -colour and sequence-space synaesthesias compared to the other groups of synaesthetes who only experienced one type or the other. More specifically, we wanted to assess whether the fact that both-synaesthetes had sequence-space synaesthesia was sufficient to show the same attentional advantages observed for sequence-synaesthetes (i.e. those without additional -colour experiences). Or, in the contrary, if having colour-synaesthesia would diminish or abolish the effects on attention for this group [both-synaesthetes]. Visual inspection of the means suggested that both-synaesthetes tended to show a similar pattern to that of colour-synaesthetes (only) in the same task for which sequence-synaesthetes showed advantages (i.e. visuo-tactile, vision task-relevant version). However, the differences were not statistically significant. Thus, future studies should properly test whether experiencing colour-synaesthesia might somewhat impair attentional advantages of people with more than one type of synaesthesia (i.e. both-synaesthetes), or if both-synaesthetes might present a specific attentional profile that is different from other synaesthetes who only have -colour or only have sequence-space types.

In addition to both-synaesthetes, while screening for participants in Study 3 we encountered several individuals who failed the -colour and sequence-space consistency tests but reported to have other types of synaesthesia that could not be assessed [with objective synaesthetic consistency tests], and who we labelled other-synaesthetes. Since this group of people did not fit our sample inclusion criteria, we removed them from the main analyses. But we also examined their performance with respect to the rest of the groups in an exploratory approach. The analyses revealed that other-synaesthetes were more efficient than non-synaesthetes at ignoring tactile distractors simultaneously presented with visual targets. That is, this group showed benefits in the same key visuo-tactile, vision task-relevant task which sequence-synaesthetes did. Interestingly, no differences were found between other-synaesthetes and colour- or sequence-synaesthetes. Future studies should specifically address attentional similarities and differences of other-synaesthetes with respect to different types of synaesthetes, but these results seem to reinforce the idea that the types of synaesthesia that synaesthetes experience might be influential at several levels of cognition and behaviour.

Synaesthetes do not only differ in terms of synaesthesia types. Importantly, during our screening interviews we observed that participants described experiencing synaesthesia in different degrees of intensity, frequency, stability, etc. Broadly speaking, these specifications could be considered different ways of defining synaesthetic strength, and one relevant question was whether this factor might have a moderating effect on attention abilities. Synaesthetic strength was operationalised in three different ways. First, we defined strength as the number of synaesthesia types experienced. Second, we defined strength as the overall degree of synaesthetic experience reported. As part of the screening process of Study 3, participants responded to an interview (Edinburgh Synaesthesia Screening Assessment; see below) that asked them to rate 'How much' each synaesthesia type applied to them choosing from a 5-point scale that ranged from 'Not at all' to 'Completely'. These ratings were then transformed into a global score (i.e. overall degree of synaesthetic experience) and used here. Third, we operationalised strength in terms of -colour and sequence-space consistency tests scores. In an exploratory approach as well, we assessed the relationship between these different measures of synaesthetic strength and participants' performance in those tasks in which we observed group differences in the main analyses (i.e. visuo-tactile tasks: vision task-irrelevant version – Study 1; and vision task-relevant version – Study 3), only within the synaesthetes subsamples and independently for colour- and sequence-synaesthetes in Study 3 given the observed differences between these two groups. The analyses showed that none of the synaesthetic strength measures and performance results in the tasks of interest were related. Therefore, these results do not suggest that variations in synaesthetic strength might moderate synaesthetes' attention abilities.

However, a different scenario was observed in terms of personality traits. In order to examine the scope of individual differences between different types of synaesthetes, Chapter III (Study 4) examined personality trait characteristics in synaesthetes with a focus on this inter-subject variability factor. Previous evidence has shown that synaesthetes have a distinct personality profile compared to non-synaesthetes, but there are inconsistencies in the literature with respect to the personality traits that differ. Most studies have focused on a particular type of synaesthesia, namely grapheme-colour synaesthesia. Thus, one possibility is that these inconsistencies are due to the (unknown or acknowledged) presence of different types of synaesthetes. To address this, we compared matched groups of colour-synaesthetes, sequence-synaesthetes, and non-synaesthetes on the Big Five personality traits (Big Five Inventory: BFI) and specific empathy (Interpersonal Reactivity Index; IRI) and positive schizotypy (Oxford-Liverpool Inventory of Feelings and Experiences; O-LIFE) questionnaires. We replicated previous findings that synaesthetes experienced higher rates of BFI Openness to Experience, IRI Fantasising (a dimension of empathy), and O-LIFE Unusual Experiences (positive schizotypy) compared to non-synaesthetes.

But this was a partial replication, as some of these differences only affected sequence-synaesthetes. In particular, this group of individuals showed higher rates of BFI Openness to Experience than non-synaesthetes *and* colour-synaesthetes. Moreover, colour- and non-synaesthetes did not differ in their scores for this trait. The same pattern was observed in relation to IRI Fantasising (i.e. higher rates for sequence-synaesthetes compared to both non- *and* colour-synaesthetes), but these group comparisons did not survive corrections. As in the attention investigations, in this study we also had additional groups of both- and other-synaesthetes. Exploratory analyses revealed that these synaesthetes showed a similar pattern to sequence-synaesthetes in relation to the BFI Openness to Experience trait, with higher rates compared to non-synaesthetes – but no significant differences compared to colour-synaesthetes in this case. Furthermore, like the sequence- and colour-synaesthetes groups, both- and other-synaesthetes also presented higher rates for the O-LIFE Unusual Experiences subscale than controls.

A secondary aim of this study was to investigate if the same synaesthetic strength variables assessed in the behavioural studies also modulated personality trait differences between synaesthetes here. Regarding sequence-synaesthetes, the results only showed positive relationships between the overall degree of synaesthetic experience reported and IRI Fantasising and O-LIFE Unusual Experiences rates, but the analyses did not survive multiple comparison corrections. On the other hand, although no differences for colour-synaesthetes were observed in terms of group comparisons, the strength analyses interestingly revealed that the greater the number of synaesthesia types and the greater the overall degree of synaesthetetic experience, the higher the rates for the O-LIFE Unusual Experiences subscale in this subgroup. Thus, whilst we did not observe modulations on attention abilities by synaesthetic strength in the Chapter II studies, this quantitative factor appears to be influential in terms of the degree of personality trait rates experienced by synaesthetes – and differently affecting different types of synaesthetes.

Considering all Chapter II and III findings, it seems that being a synaesthete, which type, and in which degree (according to different operationalisations or definitions), are all different variables that might need to be taken into account when investigating cognitive functions like attention in synaesthetes or other aspects concerning this population such as personality traits. However, we cannot ignore the fact that all these exploratory analyses present small subsamples and thus might lack power. In addition, although we have tried to apply all the due statistical corrections, all these analyses increase error rates due to multiple comparison issues. Because of all this, we want to highlight that is necessary that the identified variables of interest are properly addressed and examined in future investigations. On another note, it is worth mentioning that we grouped synaesthetes by consistency-verified types of synaesthesias (i.e. -colour and sequence-space, in particular), but this does not mean that most of our participants experienced other types of synaesthesias. In fact, we encountered high diversity both in terms of number and types of 'extra' synaesthetic experiences (see the Methods sections of each study). As a first approach to the topic of synaesthetic inter-variability, we did not address these complex differences, but given the results observed, this might have to be considered in future studies.

At the beginning of this thesis, we asked ourselves if synaesthetes were different from one another. Across the different studies conducted here, results seem to clearly suggest that synaesthetes do *not* appear to be a homogenous category of individuals. As a matter of fact, our findings point to a strong heterogeneity within synaesthetes which calls into question whether we should be considering synaesthesia a single condition. Some authors argue that, despite the causes of (developmental) synaesthesia are genetic, the specific forms that arise are possibly determined by additional factors, stochastic or random variation in brain development being its most likely contributor (Newell & Mitchell, 2016). In particular, different genetic mutations would differently affect the development of brain connectivity giving rise to a variety of synaesthetic phenotypes, from affecting one region more than another causing different types of synaesthesia to specific cross-connectivity patterns generating projector vs. associator differentiations. In addition, Newell and Mitchell (2016) propose that explicit learning and semantic memory play an important role in determining and consolidating the final inducer-concurrent associations of synaesthetes, placing thus an even greater influence on individual differences' variability and in line with our observations in the present work and those of other investigations (e.g. Bankieris & Aslin, 2016a; 2016b; Bankieris, Qian, & Aslin, 2019; Bock, 2018; Shriki, Sadeh, & Ward, 2016; Watson, Akins, Spiker, Crawford, & Enns, 2014; Whitthoft & Winawer, 2013).

For all these reasons, the need to properly screen synaesthetes from non-synaesthetes and to classify different types of synaesthetes is crucial for synaesthesia research. The last chapter of this thesis addressed this. Chapter IV (Study 5) presented the development and validation of a synaesthesia screening questionnaire, the Edinburgh Synaesthesia Screening Assessment (ESSA). Even though synaesthetic consistency tests are considered the 'gold-standard' of synaesthesia measurement, they present their own problems, the stability and reliability of synaesthetic consistency being recently questioned (see section 4.1) and, also importantly, the fact that only a few synaesthesia types can be assessed with these tools. Additional measurements such as behavioural tasks like the synaesthetic Stroop or the embedded shapes task (see section 2.1.3) might be used in convergence. But, perhaps, we should ask ourselves whether, or to what extent, synaesthesia can be measured. For these reasons, together with phenomenological approaches, self-report questionnaires and interviews are not only good instruments to screen individuals while addressing synaesthetic variability at the same time, but their importance in synaesthetic assessment might be key. However, no systematic or standardised such assessment exist to date. To that end, we developed the ESSA, a self-report questionnaire that aims to cover an exhaustive range of synaesthesia types (108) and to assess both synaesthetes and non-synaesthetes. The ESSA is organised in thirteen different categories of synaesthesia types that share the same induced concurrents and/or common characteristics (e.g. coloured sounds or mirror synaesthesias) and asks responders to rate how much each different synaesthetic experience applies to them having to choose from a 5-point Likert scale ranging from 'Not at all' to 'Completely' (5-point Likert scale). Sensitivity and specificity analyses were carried out on calculated ESSA global scores obtained from a sample of over 150 participants (including synaesthetes and non-synaesthetes) who also completed synaesthetic consistency tests for -colour and sequence-space synaesthesias. Synaesthetes obtained significantly higher scores than non-synaesthetes in the ESSA, and the analyses showed acceptable rates of sensitivity and specificity (±85.5 and ±75.8, respectively) for the questionnaire. These results were internally and externally validated (in a new sample of 275 participants) yielding some modest values.

In sum, converging evidence suggests that synaesthesia is a complex phenomenon and that synaesthetes are a diverse group of individuals. Importantly, the individual differences approach adopted here has also highlighted the fact that variations in the types of synaesthesias presented or in the degree of synaesthetic strength experienced do not only seem to be of interest, but they might be variables necessary to take into account in order to understand other synaesthetic processes or characteristics. This has wider implications for the synaesthetic research area, as it suggests that grapheme-colour synaesthetes, predominantly assessed in synaesthesia studies, might not be representative of all synaesthetes. These observations might also, at least in part, explain contrasting results reported in the literature. We recommend that future investigations focus on particular synaesthesia types or that synaesthetetic individual differences are controlled for in the study designs. In addition, we believe that future studies should directly address differences between synaesthetes from qualitative (i.e. types of synaesthesias) and quantitative (i.e. degree of synaesthetic strength) points of view, as well as examine of how these two factors interrelate. Finally, all the synaesthetes assessed here were associators and experienced -visual concurrents; future researches exploring differences in other types of synaesthetes not examined here could also be relevant contributions to the area.

References

Afra, P., Funke, M., & Matsuo, F. (2009). Acquired auditory-visual synesthesia: A window to early cross-modal sensory interactions. *Psychology Research and Behavior Management*, *2*, 31.

Alsius, A., Möttönen, R., Sams, M. E., Soto-Faraco, S., & Tiippana, K. (2014). Effect of attentional load on audiovisual speech perception: evidence from ERPs. *Frontiers in psychology*, *5*, 727.

Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Experimental Brain Research*, *183*(3), 399-404.

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*(9), 839-843.

Aluja, A., Rossier, J., & Zuckerman, M. (2007). Equivalence of paper and pencil vs Internet forms of the ZKPQ-50-CC in Spanish and French samples. *Personality and Individual Differences*, *43*(8), 2022-2032.

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*(3), 301-308.

Armel, K. C., & Ramachandran, V. S. (1999). Acquired synesthesia in retinitis pigmentosa. *Neurocase*, *5*(4), 293-296.

Armstrong, R. A. (2017). Recommendations for analysis of repeated-measures designs: testing and correcting for sphericity and use of manova and mixed model analysis. *Ophthalmic and Physiological Optics*, *37*(5), 585-593.

Arnold, D. H., Wegener, S. V., Brown, F., & Mattingley, J. B. (2012). Precision of synesthetic color matching resembles that for recollected colors rather than physical colors. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), 1078.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends in cognitive sciences*, *18*(4), 177-185.

Asher, J. E., & Carmichael, D. A. (2013). Genetics and Inheritance of Synaesthesia. In J. Simner & E. M. (Eds.), *The Oxford Handbook of Synesthesia* (pp. 23-45). Oxford, UK: Oxford University Press.

Asher, J., Aitken, M. R. F., Farooqi, N., Kurmani, S., & Baron-Cohen, S. (2006). Diagnosing and phenotyping visual synaesthesia: A preliminary evaluation of the revised test of genuineness (TOG-R). Cortex, *42*, 137–146.

Baart, M., Stekelenburg, J. J., & Vroomen, J. (2014). Electrophysiological evidence for speech-specific audiovisual integration. *Neuropsychologia*, *53*, 115-121.

Bair, W. N., Kiemel, T., Jeka, J. J., & Clark, J. E. (2012). Development of multisensory reweighting is impaired for quiet stance control in children with developmental coordination disorder (DCD). *PloS One* 7(7), e40932.

Balduzzi, S., Rücker, G., Schwarzer, G. (2019). How to perform a meta-analysis with R: a practical tutorial. *Evidence-Based Mental Health*.

Banissy, M. J., & Ward, J. (2007). Mirror-touch synaesthesia is linked with empathy. *Nature Neuroscience, 10*(7), 815.

Banissy, M. J., Cassell, J. E., Fitzpatrick, S., Ward, J., Walsh, V. X., & Muggleton, N. G. (2012). Increased positive and disorganised schizotypy in synaesthetes who experience colour from letters and tones. *cortex*, *48*(8), 1085-1087.

Banissy, M. J., Holle, H., Cassell, J., Annett, L., Tsakanikos, E., Walsh, V., Spiller, M. J., & Ward, J. (2013a). Personality traits in people with synaesthesia: Do synaesthetes have an atypical personality profile? *Personality and Individual Differences*, *54*(7), 828-831.

Banissy, M. J., Tester, V., Muggleton, N. G., Janik, A. B., Davenport, A., Franklin, A., Walsh, V., & Ward, J. (2013b). Synesthesia for color is linked to improved color perception but reduced motion perception. *Psychological science*, *24*(12), 2390-2397.

Banissy, M. J., Walsh, V., & Ward, J. (2009). Enhanced sensory perception in synaesthesia. *Experimental brain research*, *196*(4), 565-571.

Bankieris, K. R., & Aslin, R. N. (2016a). Explicit associative learning and memory in synesthetes and nonsynesthetes. *i-Perception*, 7(5), 2041669516658488.

Bankieris, K. R., & Aslin, R. N. (2016b). Implicit associative learning in synesthetes and nonsynesthetes. *Psychonomic Bulletin & Review*, 1-9.

Bankieris, K. R., & Simner, J. (2015). What is the link between synaesthesia and sound symbolism? *Cognition, 136*, 186-195.

Bankieris, K. R., Qian, T., & Aslin, R. N. (2019). Synesthetes perseverate in implicit learning: Evidence from a non-stationary statistical learning task. *Quarterly Journal of Experimental Psychology*, *72*(7), 1771-1779.

Bargary, G., Barnett, K. J., Mitchell, K. J., & Newell, F. N. (2009). Colored-speech synaesthesia is triggered by multisensory, not unisensory, perception. *Psychological Science*, *20*(5), 529-533.

Barnett, K. J., & Newell, F. N. (2008). Synaesthesia is associated with enhanced, self-rated visual imagery. *Consciousness and cognition*, *17*(3), 1032-1039.

Barnett, K. J., Foxe, J. J., Molholm, S., Kelly, S. P., Shalgi, S., Mitchell, K. J., & Newell, F. N. (2008). Differences in early sensory-perceptual processing in synesthesia: a visual evoked potential study. *Neuroimage*, *43*(3), 605-613.

Baron-Cohen, S., Burt, L., Laittan-Smith, F., Harrison, J. E., & Bolton, P. (1996). Synaesthesia: Prevalence and familiarity. *Perception, 25*, 1073-1079.

Baron-Cohen, S., Harrison, J., Goldstein, L. H., & Wyke, M. (1993). Coloured speech perception: Is synaesthesia what happens when modularity breaks down? *Perception, 22*(4), 419–426.

Baron-Cohen, S., Johnson, D., Asher, J., Wheelwright, S., Fisher, S. E., Gregersen, P. K., & Allison, C. (2013). Is synaesthesia more common in autism? *Molecular autism*, *4*(1), 40.

Baron-Cohen, S., Robson, E., Lai, M. C., & Allison, C. (2016). Mirror-touch synaesthesia is not associated with heightened empathy, and can occur with autism. *PloS one*, *11*(8), e0160543.

Baron-Cohen, S., Wyke, M., & Binnie, C. (1987). Hearing words and seeing colours: An experimental investigation of a case of synaesthesia. *Perception, 16*, 761-767.

Barton, K. (2018, August 15). *MuMIn: Multi-Model Inference. R package version 1.42.1*. Retrieved from https://CRAN.R-project.org/package=MuMIn

Bates, D., & Maechler, M. (2018, May 15). *Matrix: Sparse and Dense Matrix Classes and Methods. R package version 1.2-14*. Retrieved from https://CRAN.R-project.org/package=Matrix

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67(*1), 1-48.

Bien, N., ten Oever, S., Goebel, R. & Sack, A. T. (2012). The sound of size: crossmodal binding in pitch-size synesthesia: a combined TMS, EEG and psychophysics study. *Neuroimage* 59(1), 663-672.

Bock, J. (2018). A deep learning model of perception in Color-Letter synesthesia. *Big Data and Cognitive Computing*, *2*(1), 8.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to Meta-Analysis*. John Wiley & Sons.

Borst, G., & Kosslyn, S. M. (2008). Visual mental imagery and visual perception: Structural equivalence revealed by scanning processes. *Memory & Cognition*, *36*(4), 849-862.

Brang, D., & Ramachandran, V. S. (2020). How do crossmodal correspondences and multisensory processes relate to synesthesia?. In *Multisensory Perception* (pp. 259-281). Academic Press.

Brang, D., Miller, L. E., McQuire, M., Ramachandran, V. S., & Coulson, S. (2013a). Enhanced mental rotation ability in time-space synesthesia. *Cognitive processing*, *14*(4), 429-434.

Brang, D., Taich, Z. J., Hillyard, S. A., Grabowecky, M., & Ramachandran, V. S. (2013b). Parietal connectivity mediates multisensory facilitation. *Neuroimage*, *78*, 396-401.

Brang, D., Teuscher, U., Ramachandran, V. S., & Coulson, S. (2010). Temporal sequences, synesthetic mappings, and cultural biases: The geography of time. *Consciousness and Cognition*, *19*(1), 311-320.

Brang, D., Williams, L. E., & Ramachandran, V. S. (2012). Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities. *Cortex, 48*(5), 630-637.

Brüggen, E., & Dholakia, U. M. (2010). Determinants of participation and response effort in web panel surveys. *Journal of Interactive Marketing*, *24*(3), 239-250.

Bujang, M. A., & Adnan, T. H. (2016). Requirements for minimum sample size for sensitivity and specificity analysis. *Journal of clinical and diagnostic research: JCDR*, *10*(10), YE01.

Burghoorn, F., Dingemanse, M., van Lier, R., & van Leeuwen, T. M. (2019). The relation between autistic traits, the degree of synaesthesia, and local/global visual perception. *Journal of autism and developmental disorders*, 1-18.

Cacace, A. T., Cousins, J. P., Parnes, S. M., McFarland, D. J., Semenoff, D., Holmes, T., Davenport, C., Stegbauer, K., & Lovely, T. J. (1999). Cutaneous-evoked tinnitus. *Audiology and Neurotology*, *4*(5), 258-268.

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1-3), 191-205.

Carlbring, P., Brunt, S., Bohman, S., Austin, D., Richards, J., Öst, L. G., & Andersson, G. (2007). Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior*, *23*(3), 1421-1434.

Carmichael, D. A., & Simner, J. (2013). The immune hypothesis of synesthesia. *Frontiers in human neuroscience*, *7*, 563.

Carmichael, D. A., Down, M. P., Shillcock, R. C., Eagleman, D. M., & Simner, J. (2015). Validating a standardised test battery for synesthesia: Does the Synesthesia Battery reliably detect synesthesia? *Consciousness and cognition*, *33*, 375-385.

Carmichael, D. A., Smees, R., Shillcock, R. C., & Simner, J. (2019). Is there a burden attached to synaesthesia? Health screening of synaesthetes in the general population. *British Journal of Psychology*, *110*(3), 530-548.

Carruthers, H. R., Miller, V., Tarrier, N., & Whorwell, P. J. (2012). Synesthesia, pseudo-synesthesia, and irritable bowel syndrome. *Digestive diseases and sciences*, *57*(6), 1629-1635.

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 367-379.

Chan, D. (2009). So why ask me? Are self-report data really that bad. *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences*, 309-336.

Chiou, R., & Rich, A. N. (2014). The role of conceptual knowledge in understanding synesthesia: evaluating contemporary findings from a “hub-and-spokes” perspective. *Frontiers in Psychology, 5*, 105.

Choi, B. C., & Pak, A. W. (2005). Peer reviewed: a catalog of biases in questionnaires. *Preventing chronic disease*, *2*(1).

Chou, R., Dana, T., & Bougatsos, C. (2009). Screening older adults for impaired visual acuity: a review of the evidence for the US Preventive Services Task Force. *Annals of internal medicine*, *151*(1), 44-58.

Chun, C. A., & Hupé, J. M. (2013). Mirror-touch and ticker tape experiences in synesthesia. *Frontiers in psychology*, *4*, 776.

Chun, C. A., & Hupé, J. M. (2016). Are synesthetes exceptional beyond their synesthetic associations? A systematic comparison of creativity, personality, cognition, and mental imagery in synesthetes and controls. *British Journal of Psychology*, *107*(3), 397-418.

Cohen-Kadosh, R., & Walsh, V. (2008). Synaesthesia and cortical connections: cause or correlation? *Trends in Neuroscience, 31*(11), 549-550.

Cohen-Kadosh, R., Henik, A., Catena, A., Walsh, V., & Fuentes, L. J. (2009). Induced cross-modal synaesthetic experience without abnormal neuronal connections. *Psychological Science*, *20*(2), 258-265.

Costa, P. T., & McCrae, R. R. (1989). *NEO five-factor inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources.

Cytowic, R. E. (1989). Synesthesia and mapping of subjective sensory dimensions. *Neurology*, *39*(6), 849-850.

Cytowic, R. E., & Eagleman, D. M. (2009). *Wednesday is Indigo Blue: Discovering the Brain of Synesthesia*. Boston, MA: MIT Press.

Cytowic, R.E. (2002). *Synesthesia: A union of the senses*. MIT press, Boston, MA, USA.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy.

Day, S. (2005). Some demographic and socio-cultural aspects of synaesthesia. In L. C. Robertson & Sagiv, N. (Eds.), *Synesthesia: Perspectives from cognitive neuroscience*. Oxford, UK: Oxford University Press.

Day, S. (2019, October 22). *Demographic aspects of synesthesia*. Retrieved from http://www.daysyn.com/Types-of-Syn.html.

De Haas, B., Kanai, R., Jalkanen, L., & Rees, G. (2012). Grey matter volume in early human visual cortex predicts proneness to the sound-induced flash illusion. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 4955-4961.

DerSimonian, R., & Laird, N. (1986). Meta-Analysis in Clinical Trials. *Controlled Clinical Trials,* 7(3), 177–88.

Dixon, M. J., & Smilek, D. (2005). The importance of individual differences in grapheme-color synesthesia. *Neuron*, *45*(6), 821-823.

Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: Projector vs. associator synaesthetes. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(3), 335-343.

Dojat, M., Pizzagalli, F., & Hupé, J. M. (2018). Magnetic resonance imaging does not reveal structural alterations in the brain of grapheme-color synesthetes. *PloS one*, *13*(4).

Dojat, M., Pizzagalli, F., & Hupé, J. M. (2018). Magnetic resonance imaging does not reveal structural alterations in the brain of grapheme-color synesthetes. *PloS one*, *13*(4), e0194422.

Dollinger, S. J., & Leong, F. T. (1993). Volunteer bias and the five-factor model. *The Journal of psychology*, *127*(1), 29-36.

Donker, T., van Straten, A., Marks, I., & Cuijpers, P. (2010). Brief self-rated screening for depression on the Internet. *Journal of affective disorders*, *122*(3), 253-259.

Donnellan, M. B., Lucas, R. E., & Cesario, J. (2015). On the association between loneliness and bathing habits: Nine replications of Bargh and Shalev (2012) Study 1. *Emotion*, *15*(1), 109.

Dresslar, F. B. (1903). Are chromaesthesias variable? A study of an individual case. *American Journal of Psychology, 14*, 632–646.

Driver, J., & Spence, C. (1998a). Attention and the crossmodal construction of space. *Trends in cognitive sciences*, *2*(7), 254-262.

Driver, J., & Spence, C. (1998b). Crossmodal attention. *Current opinion in neurobiology*, *8*(2), 245-253.

Eagleman, D. M. (2009). The objectification of overlearned sequences: A new view of spatial sequence synesthesia. *Cortex, 45*(10), 1266–1277. http://dx.doi.org/10.1016/j.cortex.2009.06.012

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007). A standardized test battery for the study of synesthesia. *Journal of neuroscience methods*, *159*(1), 139-145.

Edquist, J., Rich, A. N., Brinkman, C., & Mattingley, J. B. (2006). Do synaesthetic colours act as unique features in visual search? *Cortex*, *42*(2), 222-231.

Eimer, M., & Schröger, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology, 35*(03), 313-327.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception &. Psychophysics, 16*(1), 143-149.

Eriksen, C. W. (1995). The flanker task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual. Cognition, 2*(2-3), 101-118.

Esterman, M., Verstynen T., Ivry R. B., & Robertson L. C. (2006). Coming unbound: disrupting automatic integration of synesthetic color and graphemes by transcranial magnetic stimulation of the right parietal lobe. *Journal of cognitive neuroscience, 18*(9), 1570-1576.

Facoetti, A., Trussardi, A. N., Ruffino, M., Lorusso, M. L., Cattaneo, C., Galli, R., Molteni, M., & Zorsi, M. (2010). Multisensory spatial attention deficits are predictive of phonological decoding skills in developmental dyslexia. *Journal of cognitive neuroscience, 22*(5), 1011-1025.

Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., & Posner, M. I. (2003). Cognitive and brain consequences of conflict. *Neuroimage, 18*(1), 42-57.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, *27*(8), 861-874.

Follmann, D. A., & Proschan, M. A. (1999). Valid Inference in Random Effects Meta-Analysis. *Biometrics, 55*(3), 732–37.

Fornazzari, L., Fischer, C. E., Ringer, L., & Schweizer, T. A. (2012). “Blue is music to my ears”: Multimodal synesthesias after a thalamic stroke. *Neurocase*, *18*(4), 318-322.

Forster, B., & Pavone, E. F. (2008). Electrophysiological correlates of crossmodal visual distractor congruency effects: evidence for response conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 65-73.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P., & Ridderinkhof, K. R. (2008). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. *Journal of Neuroscience*, *28*(39), 9790-9796.

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, *203*(2), 381-389.

Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression* (2nd ed.). Thousand Oaks, CA: Sage.

Fox, J., Weisberg, S., & Price, B. (2019, May 2). *carData: Companion to Applied Regression Data Sets. R packag e version 3.0-1*. Retrieved from https://CRAN.R-project.org/package=carData

Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2000). Multisensory auditory–somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cognitive Brain Research*, *10*(1-2), 77-83.

Galton, F. (1880). Visualised numerals. *Nature, 21*, 252-256.

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research*, *20*(2), 226-241.

Gardiner, J. C., Luo, Z., & Roman, L. A. (2009). Fixed effects, random effects and GEE: what are the differences?. *Statistics in medicine*, *28*(2), 221-239.

Gertner, L., Arend, I., & Henik, A. (2012). Effects of non-symbolic numerical information suggest the existence of magnitude–space synesthesia. *Cognitive processing*, *13*(1), 179-183.

Gheri, C., Chopping, S., & Morgan, M. J. (2008). Synaesthetic colours do not camouflage form in visual search. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1636), 841-846.

Ginsberg, L. (1923). A case of synaesthesia. *American Journal of Psychology, 34*, 582-589.

Glas, A. S., Lijmer, J. G., Prins, M. H., Bonsel, G. J., & Bossuyt, P. M. (2003). The diagnostic odds ratio: a single indicator of test performance. *Journal of clinical epidemiology*, *56*(11), 1129-1135.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American psychologist*, *48*(1), 26.

Goller, A. I., Richards, K., Novak, S., & Ward, J. (2013). Mirror-touch synaesthesia in the phantom limbs of amputees. *Cortex*, *49*(1), 243-251.

Green, J. A., & Goswami, U. (2008). Synesthesia and number cognition in children. *Cognition*, *106*(1), 463-473.

Gross, V. C., Neargarder, S., Caldwell-Harris, C. L., & Cronin-Golomb, A. (2011). Superior encoding enhances recall in color-graphemic synesthesia. *Perception*, *40*(2), 196-208.

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends in cognitive sciences*, *5*(1), 36-41.

Hajian-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian journal of internal medicine*, *4*(2), 627.

Hale, J., Thompson, J. M., Morgan, H. M., Cappelletti, M., & Cohen-Kadosh, R. (2014). Better together? The cognitive advantages of synaesthesia for time, numbers, and space. *Cognitive neuropsychology*, *31*(7-8), 545-564.

Hänggi, J., Beeli, G., Oechslin, M. S., & Jäncke, L. (2008). The multiple synaesthete ES—Neuroanatomical basis of interval-taste and tone-colour synaesthesia. *Neuroimage*, *43*(2), 192-203.

Harrer M, Cuijpers P, Furukawa T, Ebert DD (2019). *dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R'*. R package version 0.0.9000, http://dmetar.protectlab.org.

Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). *Doing Meta-Analysis in R: A Hands-on Guide*. DOI: 10.5281/zenodo.2551803.

Hartung, J. (1999). An Alternative Method for Meta-Analysis. *Biometrical Journal: Journal of Mathematical Methods in Biosciences, 41*(8), 901–16.

Hartung, J., & Knapp, G. (2001). A Refined Method for the Meta-Analysis of Controlled Clinical Trials with Binary Outcome. *Statistics in Medicine, 20*(24), 3875–89.

Havlik, A. M., Carmichael, D. A., & Simner, J. (2015). Do sequence-space synaesthetes have better spatial imagery skills? Yes, but there are individual differences. *Cognitive processing*, *16*(3), 245-253.

Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. *Journal of Educational Statistics, 6*(2), 107–28.

Higgins, J. P. T, & Thompson, S. G. (2002). Quantifying Heterogeneity in a Meta-Analysis. *Statistics in Medicine, 21*(11), 1539–58.

Holle, H., Banissy, M., Wright, T. D., Bowling, N., & Ward, J. (2011). "That's not a real body": Identifying stimulus qualities that modulate synaesthetic experiences of touch. *Consciousness and Cognition, 20*, 720–726.

Holmes, N. P. (2012). Does tool use extend peripersonal space? A review and re-analysis. *Experimental brain research*, *218*(2), 273-282.

Holmes, N. P., Calvert, G. A., & Spence, C. (2004). Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neuroscience letters*, *372*(1-2), 62-67.

Holmes, N. P., Calvert, G. A., & Spence, C. (2007). Tool use changes multisensory interactions in seconds: evidence from the crossmodal congruency task. *Experimental Brain Research*, *183*(4), 465-476.

Holmes, N. P., Sanabria, D., Calvert, G. A., & Spence, C. (2006). Multisensory interactions follow the hands across the midline: Evidence from a non-spatial visual-tactile congruency task. *Brain research, 1077*(1), 108-115.

Hossain, S. R., Simner, J., & Ipser, A. (2018). Personality predicts the vibrancy of colour imagery: The case of synaesthesia. *Cortex, 105*, 74-82.

Hötting, K., Rösler, F., & Röder, B. (2003). Crossmodal and intermodal attention modulate event-related brain potentials to tactile and auditory stimuli. *Experimental Brain Research, 148*(1), 26-37.

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005). Individual differences among grapheme-color synesthetes: brain-behavior correlations. *Neuron*, *45*(6), 975-985.

Hubbard, E. M., Ranzini, M., Piazza, M., & Dehaene, S. (2009). What information is critical to elicit interference in number-form synaesthesia? *Cortex*, *45*(10), 1200-1216.

Hughes, J.E.A., Simner, J., Baron-Cohen, S., Treffert, D. A., & Ward, J. (2017). Is synaesthesia more prevalent in autism spectrum conditions? Only where there is prodigious talent. *Multisensory Research, 30*(3-5), 391-408.

Hupé, J. M., & Dojat, M. (2015). A critical review of the neuroimaging literature on synesthesia. *Frontiers in human neuroscience*, *9*, 103.

IBM Corporation. (2016). *IBM SPSS Statistics for Windows* (Version 24.0) [Computer Software].

Igarashi, Y., Kitagawa, N., Spence, C., & Ichihara, S. (2007). Assessing the influence of schematic drawings of body parts on tactile discrimination performance using the crossmodal congruency task. *Acta Psychologica*, *124*(2), 190-208.

IntHout, J., Ioannidis, J. P. A., & Borm, G. F. (2014). The Hartung-Knapp-Sidik-Jonkman Method for Random Effects Meta-Analysis Is Straightforward and Considerably Outperforms the Standard Dersimonian-Laird Method. *BMC Medical Research Methodology. 14*(1).

Ipser, A., Ward, J., & Simner, J. (2019). The MULTISENSE Test of Lexical–Gustatory Synaesthesia: An automated online diagnostic. *Behavior research methods*, 1-17.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (Eds.). (2013). *An introduction to statistical learning* (Vol. 103). New York: Springer.

Jamovi Project (2018). *Jamovi* (Version 0.9) [Computer Software].

Jäncke, L., Beeli, G., Eulig, C., & Hänggi, J. (2009). The neuroanatomy of grapheme–color synesthesia. *European Journal of Neuroscience*, *29*(6), 1287-1293.

Janik McErlean, A. B., & Banissy, M. J. (2016). Examining the relationship between schizotypy and self-reported visual imagery vividness in grapheme-color synaesthesia. *Frontiers in psychology*, *7*.

Jansari, A. S., Spiller, M. J., & Redfern, S. (2006). Number synaesthesia: When hearing "four plus five" looks like gold. *Cortex*, *42*(2), 253-258.

JASP Team (2018). *JASP* (Version 0.9) [Computer software].

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). New York: Oxford University Press.

Jewanski, J., Simner, J., Day, S. A., & Ward, J. (2011). The development of a scientific understanding of synesthesia from early case studies (1849–1873). *Journal of the History of the Neurosciences*, *20*(4), 284-305.

Jobe, J. B. (2003). Cognitive psychology and self-reports: Models and methods. *Quality of Life Research*, *12*(3), 219-227.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research, 2*(1999), 102-138.

Johnson, D., Allison, C., & Baron-Cohen, Simon. (2013). The prevalence of synesthesia: The Consistency Revolution. In J. Simner & E. M. Hubbard (Eds), *Oxford Handbook of Synesthesia*. Oxford, UK: Oxford University Press.

Jonas, C. N., & Hibbard, P. B. (2015). Migraine in synesthetes and nonsynesthetes: A prevalence study. *Perception*, *44*(10), 1179-1202.

Jonas, C. N., & Price, M. C. (2014). Not all synesthetes are alike: spatial vs. visual dimensions of sequence-space synesthesia. *Frontiers in psychology*, *5*, 1171.

Jürgens, T. P., Schulte, L. H., & May, A. (2014). Migraine trait symptoms in migraine with and without aura. *Neurology*, *82*(16), 1416-1424.

Kamke, M. R., Vieth, H. E., Cottrell, D., & Mattingley, J. B. (2012). Parietal disruption alters audiovisual binding in the sound-induced flash illusion. *NeuroImage*, *62*(3), 1334-1341.

Kay, C. L., Carmichael, D. A., Ruffell, H. E., & Simner, J. (2015). Colour fluctuations in grapheme-colour synaesthesia: The effect of clinical and non-clinical mood changes. *British Journal of Psychology*, *106*(3), 487-504.

Kennis, M., Rademaker, A. R., & Geuze, E. (2013). Neural correlates of personality: an integrative review. *Neuroscience & Biobehavioral Reviews*, *37*(1), 73-95.

Klein Richard, A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bahník, Š., Bernstein, M. J., ... & Cemalcilar, Z. (2014). Investigating Variation in Replicability: A ‘‘Many Labs’’Replication Project. *Social Psychology*, *45*(3), 142-152.

Knowland, V. C., Mercure, E., Karmiloff-Smith, A., Dick, F., & Thomas, M. S. (2014). Audio-visual speech perception: A developmental ERP investigation. *Developmental Science*, *17*(1), 110-124.

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta psychologica*, *134*(3), 372-384.

Köhler, W. (1929). *Gestalt Psychology*. New York: Liveright.

Kosslyn, S. M. (1994). *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.

Krypotos, A. M., Blanken, T. F., Arnaudova, I., Matzke, D., & Beckers, T. (2017). A primer on Bayesian analysis for experimental psychopathologists. *Journal of experimental psychopathology*, *8*(2), 140-157.

Kuhn, M. (2019, April 27). *caret: Classification and Regression Training. R package version 6.0-84*. Retrieved fro m https://CRAN.R-project.org/package=caret

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*(Vol. 26). New York: Springer.

Kusnir, F., & Thut, G. (2012). Formation of automatic letter–colour associations in non-synaesthetes through likelihood manipulation of letter–colour pairings. *Neuropsychologia,* 50(14), 3641-3652.

Kuznetsova A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26.

Laeng, B. (2009). Searching through synaesthetic colors. *Attention, Perception, & Psychophysics*, *71*(7), 1461-1467.

Laeng, B., Hugdahl, K., & Specht, K. (2011). The neural correlate of colour distances revealed with competing synaesthetic and real colours. *Cortex*, *47*(3), 320-331.

Laeng, B., Svartdal, F., & Oelmann, H. (2004). Does color synesthesia pose a paradox for early-selection theories of attention? *Psychological Science*, *15*(4), 277-281.

LeBel, E. (2015). A new replication norm for psychology. *Collabra: Psychology*, *1*(1).

Lee, M. D., & Wagenmakers. E.-J. (2005). Bayesian statistical inference in psychology: Comment on Trafimow (2003). *Psychological Review, 112*(3), 662-668.

Lilienfeld, S. O., & Waldman, I. D. (Eds.). (2017). *Psychological science under scrutiny: Recent challenges and proposed solutions*. John Wiley & Sons.

Long, J. A. (2019, April 8). *jtools: Analysis and Presentation of Social Scientific Data. R package version 1.1.1*. Re trieved from https://cran.r-project.org/package=jtools

Lönnqvist, J. E., Paunonen, S., Verkasalo, M., Leikas, S., Tuulio-Henriksson, A., & Lönnqvist, J. (2007). Personality characteristics of research volunteers. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(8), 1017-1030.

Lunke, K., & Meier, B. (2018). New insights into mechanisms of enhanced synaesthetic memory: Benefits are synaesthesia-type-specific. *PloS one*, *13*(9), e0203055.

Ma, Y., Mazumdar, M., & Memtsoudis, S. G. (2012). Beyond repeated-measures analysis of variance: advanced statistical methods for the analysis of longitudinal data in anesthesia research. *Regional Anesthesia & Pain Medicine*, *37*(1), 99-105.

Macaluso, E., Frith, C.D., & Driver, J. (2002). Crossmodal spatial influences of touch on extrastriate visual areas take current gaze direction into account. *Neuron*, *34*(4),647-658.

Macaluso, E., Noppeney, U., Talsma, D., Vercillo, T., Hartcher-O'Brien, J., & Adam, R. (2016). The curious incident of attention in multisensory integration: bottom-up vs. top-down. *Multisensory Research*, *29*(6-7), 557-583.

MacLeod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology* *24*(5), 637.

Makambi, K. H. (2004). The Effect of the Heterogeneity Variance Estimator on Some Tests of Treatment Efficacy. *Journal of Biopharmaceutical Statistics, 14*(2), 439–49.

Mann, H., Korzenko, J., Carriere, J. S., & Dixon, M. J. (2009). Time–space synaesthesia–A cognitive advantage? *Consciousness and cognition*, *18*(3), 619-627.

Maravita, A., Spence, C., & Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Current biology*, *13*(13), R531-R539.

Maravita, A., Spence, C., Kennett, S., & Driver, J. (2002). Tool-use changes multimodal spatial interactions between vision and touch in normal humans. *Cognition*, *83*(2), B25-B34.

Marcus, B., & Schütz, A. (2005). Who are the people reluctant to participate in research? Personality correlates of four different types of nonresponse as inferred from self-and observer ratings. *Journal of personality*, *73*(4), 959-984.

Marini, F., Chelazzi, L., & Maravita, A. (2013). The costly filtering of potential distraction: Evidence for a supramodal mechanism. *Journal of Experimental Psychology: General*, *142*(3), 906.

Marini, F., Romano, D., & Maravita, A. (2017). The contribution of response conflict, multisensory integration, and body-mediated attention to the crossmodal congruency effect. *Experimental brain research*, *235*(3), 873-887.

Marini, F., Tagliabue, C. F., Sposito, A. V., Hernandez-Arieta, A., Brugger, P., Estévez, N., & Maravita, A. (2014). Crossmodal representation of a functional robotic hand arises after extensive training in healthy participants. *Neuropsychologia*, *53*, 178-186.

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *The American journal of psychology*, 173-188.

Mas-Casadesús, A. (2019). Is there a synaesthetic personality profile? Characterising synaesthetes' singularity. *PsyPAG Quarterly, 111*, 25-27.

Mas-Casadesús, A., & Gherri, E. (2017). Ignoring irrelevant information: Enhanced intermodal attention in synaesthetes. *Multisensory Research, 30*(3-5), 253-277.

Mason, O., & Claridge, G. (2006). The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE): further description and extended norms. *Schizophrenia research*, *82*(2-3), 203-211

Mattingley, J. B., Payne, J. M., & Rich, A. N. (2006). Attentional load attenuates synaesthetic priming effects in grapheme-colour synaesthesia. *Cortex, 42*(2), 213-221.

Mattingley, J. B., Rich, A. N., Yelland, G., & Bradshaw, J. L. (2001). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature*, *410*(6828), 580-582.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology, 52*(1), 81.

McFeely, J. W., Antonelli, P. J., Rodríguez, F. J., & Holmes, A. E. (1998). Somatosensory phenomena after multichannel cochlear implantation in prelingually deaf adults. *The American journal of otology*, *19*(4), 467-471.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica*, *22*(3), 276-282.

Meier, B., & Rothen, N. (2007). When conditioned responses "fire back": bidirectional cross-activation creates learning opportunities in synesthesia. *Neuroscience*, *147*(3), 569-572.

Meier, B., Rothen, N., & Walter, S. (2014). Developmental aspects of synaesthesia across the adult lifespan. *Frontiers in human neuroscience*, *8*, 129.

Menouti, K., Akiva-Kabiri, L., Banissy, M. J., & Stewart, L. (2015). Timbre-colour synaesthesia: exploring the consistency of associations based on timbre. *Cortex*, *63*, 1-3.

Merat, N., Spence, C., Lloyd, D. M., Withington, D. J., & McGlone, F. (1999). Audiotactile links in focused and divided spatial attention. *Society for Neuroscience Abstracts, 25*.

Milán, E., Iborra, O., de Córdoba, M., Juárez-Ramos, V., Artacho, M. R., & Rubio, J. L. (2013). The Kiki-Bouba effect A case of personification and ideaesthesia. *Journal of Consciousness Studies, 20*(1-2), 84-102.

Mishra, J., Martínez, A., & Hillyard, S. A. (2008). Cortical processes underlying sound-induced flash fusion. *Brain research*, *1242*, 102-115.

Mishra, J., Martínez, A., & Hillyard, S. A. (2010). Effect of attention on early cortical processes associated with the sound-induced extra flash illusion. *Journal of cognitive neuroscience*, *22*(8), 1714-1729.

Morís Fernández, L., Macaluso, E., & Soto-Faraco, S. (2017). Audiovisual integration as conflict resolution: The conflict of the McGurk illusion. *Human brain mapping*, *38*(11), 5691-5705.

Morís Fernández, L., Torralba, M., & Soto-Faraco, S. (2018). Theta oscillations reflect conflict processing in the perception of the McGurk illusion. *European Journal of Neuroscience*, *48*(7), 2630-2641.

Morís Fernández, L., Visser, M., Ventura-Campos, N. Ávila, C., & Soto-Faraco, S. (2015). Top-down attention regulates the neural expression of audiovisual integration. *NeuroImage, 119*, 272-285.

Muggleton, N., Tsakanikos, E., Walsh, V., & Ward, J. (2007). Disruption of synaesthesia following TMS of the right posterior parietal cortex. *Neuropsychologia*, *45*(7), 1582-1585.

MULTISENSE Research Project (2019). MULTISENSE Adaptable Synaesthesia Toolkit [Computer Software].

Neufeld, J., Roy, M., Zapf, A., Sinke, C., Emrich, H. M., Prox-Vagedes, V., Wolfang, D., & Zedler, M. (2013). Is synesthesia more common in patients with Asperger syndrome? *Frontiers in human neuroscience*, *7*, 847.

Neufeld, J., Sinke, C., Zedler, M., Emrich, H. M., & Szycik, G. R. (2012). Reduced audio–visual integration in synaesthetes indicated by the double-flash illusion. *Brain research*, *1473*, 78-86.

Newell, F. N. (2013). Synesthesia, meaning, and multilingual speakers. In J. Simner & E. M. (Eds.), *The Oxford Handbook of Synesthesia* (pp. 181-204). Oxford, UK: Oxford University Press.

Newell, F. N., & Mitchell, K. J. (2016). Multisensory integration and cross-modal learning in synaesthesia: A unifying model. *Neuropsychologia*, *88*, 140-150.

Niccolai, V., Jennes, J., Stoerig, P., & van Leeuwen, T. M. (2012). Modality and variability of synesthetic experience. *The American journal of psychology*, *125*(1), 81-94.

Nijboer, T. C., & van der Stigchel, S. (2009). Is attention essential for inducing synesthetic colors? Evidence from oculomotor distractors. *Journal of vision*, *9*(6), 21-21.

Novich, S., Cheng, S., & Eagleman, D. M. (2011). Is synaesthesia one condition or many? A large-scale analysis reveals subgroups. *Journal of Neuropsychology*, *5*(2), 353-371.

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C., Parslow, D. M., Morgan, M. J., Bullmore, E. T., Baron-Cohen, S., & Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nature neuroscience*, *5*(4), 371.

O'Dowd, A., Cooney, S. M., McGovern, D. P, & Newell, F. N. (2019). Do synaesthesia and mental imagery tap into similar cross-modal processes? *Philosophical Transactions of the Royal Society B, 374*(1787), 20180359..

Olfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

Palmeri, T. J., Blake, R., Marois, R., Flanery, M. A., & Whetsell, W. (2002). The perceptual reality of synesthetic colors. *Proceedings of the National Academy of Sciences*, *99*(6), 4127-4131.

Pasalar, S., Ro, T., & Beauchamp, M. S. (2010). TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *European Journal of Neuroscience*, *31*(10), 1783-1790.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence?. *Perspectives on Psychological Science*, *7*(6), 528-530.

Paulesu, E., Harrison, J., Baron-Cohen, S., Watson, J. D., Goldstein, L., Heather, J., Frackowiak, R. S. J., & Frith, C. D. (1995). The physiology of coloured hearing A PET activation study of colour-word synaesthesia. *Brain*, *118*(3), 661-676.

Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological science*, *11*(5), 353-359.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2019, August 1). *nlme: Linear and Nonlinear Mix ed Effects Models. R package version 3.1-137*. Retrieved from https://CRAN.R-project.org/package=nlme

Poliakoff, E., Ashworth, S., Lowe, C., & Spence, C. (2006). Vision and touch in ageing: crossmodal selective attention and visuotactile spatial interactions. *Neuropsychologia*, *44*(4), 507-517.

Poole, D., Couth, S., Gowen, E., Warren, P. A., & Poliakoff, E. (2015). Adapting the crossmodal congruency task for measuring the limits of visual–tactile interactions within and between groups. *Multisensory research*, *28*(3-4), 227-244.

Preckel, F., & Thiemann, H. (2003). Online- vs. paper-pencil version of a high potential intelligence test. *Swiss Journal of Psychology / Schweizerische Zeitschrift für Psychologie / Revue Suisse de Psychologie, 62*(2), 131-138.

Price, M. C. (2009). Spatial forms and mental imagery. *Cortex*, *45*(10), 1229-1245.-

Price, M. C. (2013). Synesthesia, imagery, and performance. In J. Simner & E. M. (Eds.), *The Oxford Handbook of Synesthesia* (pp. 728-760). Oxford, UK: Oxford University Press.

Price, M. C., & Mentzoni, R. A. (2008). Where is January? The month-SNARC effect in sequence-form synaesthetes. *Cortex*, *44*(7), 890-907.

Pritchard, J., Rothen, N., Coolbear, D., & Ward, J. (2013). Enhanced associative memory for colour (but not shape or location) in synaesthesia. *Cognition*, *127*(2), 230-234.

R Core Team (2018). R 3.5.1 [Computer software].

Rader, C. M., & Tellegen, A. (1987). An investigation of synesthesia. *Journal of Personality and Social Psychology*, *52*(5), 981.

Radvansky, G. A., Gibson, B. S., & McNerney, M. (2011). Synesthesia and memory: Color congruency, von Restorff, and false memory effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 219.

Ramachandran, V. S., & Hubbard, E. M. (2001a). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1470), 979-983.

Ramachandran, V. S., & Hubbard, E. M. (2001b). Synaesthesia: a window into perception, thought and language. *Journal of consciousness studies*, *8*(12), 3-34.

Ramachandran, V. S., Rogers-Ramachandran, D., & Cobb, S. (1995). Touching the phantom limb. *Nature, 377*, 489–490.

Raslich, M. A., Markert, R. J., & Stutes, S. A. (2007). Selecting and interpreting diagnostic tests. *Biochemia medica: Biochemia medica*, *17*(2), 151-161.

Riccelli, R., Toschi, N., Nigro, S., Terracciano, A., & Passamonti, L. (2017). Surface-based morphometry reveals the neuroanatomical basis of the five-factor model of personality. *Social Cognitive and Affective Neuroscience*, *12*(4), 671-684.

Rich, A. N., & Mattingley, J. B. (2005). Can attention modulate color-graphemic synesthesia? In L. C. Robertson & N. Sagiv (Eds.), *Synesthesia: Perspectives from Cognitive Neuroscience* (pp. 108–123). Oxford, UK: Oxford University Press.

Rich, A. N., & Mattingley, J. B. (2010). Out of sight, out of mind: The attentional blink can eliminate synaesthetic colours. *Cognition, 114*, 320–328.

Rich, A. N., & Mattingley, J. B. (2013). The role of attention in synaesthesia. In J. Simner & E. M. (Eds.), *The Oxford Handbook of Synesthesia* (pp. 265-282). Oxford, UK: Oxford University Press.

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: implications for the role of early experience in lexical-colour associations. *Cognition*, *98*(1), 53-84.

Ridderinkhof, K. R., Band, G. P., & Logan, G. D. (1999). A study of adaptive behavior: Effects of age and irrelevant information on the ability to inhibit one's actions. *Acta psychologica*, *101*(2-3), 315-337.

Riva, G., Teruzzi, T., & Anolli, L. (2003). The use of the internet in psychological research: Comparison of online and offline questionnaires. *CyberPsychology & Behavior*, *6*(1), 73-80.

Rizza, A., & Price, M. C. (2012). Do sequence-space synaesthetes have better spatial imagery skills? Maybe not. *Cognitive processing*, *13*(1), 299-303.

Ro, T., Farnè, A., Johnson, R. M., Wedeen, V., Chu, Z., Wang, Z. J., Hunter, J. V., & Beauchamp, M. S. (2007). Feeling sounds after a thalamic lesion. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *62*(5), 433-441.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sánchez, J.-C., & Mülle, M. (2011). pROC: An open-sou rce package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77.

Rothen, N., & Meier, B. (2009). Do synesthetes have a general advantage in visual search and episodic memory? A case for group studies. *PLoS One*, *4*(4), e5037.

Rothen, N., & Meier, B. (2010). Grapheme–colour synaesthesia yields an ordinary rather than extraordinary memory advantage: evidence from a group study. *Memory*, *18*(3), 258-264.

Rothen, N., Jünemann, K., Mealor, A. D., Burckhardt, V., & Ward, J. (2016). The sensitivity and specificity of a diagnostic test of sequence-space synaesthesia. *Behavior Research Methods, 48*(4), 1476–1481.

Rothen, N., Meier, B., & Ward, J. (2012). Enhanced memory ability: insights from synaesthesia. *Neuroscience & Biobehavioral Reviews*, *36*(8), 1952-1963.

Rothen, N., Nyffeler, T., von Wartburg, R., Müri, R., & Meier, B. (2010). Parieto-occipital suppression eliminates implicit bidirectionality in grapheme-colour synaesthesia. *Neuropsychologia*, *48*(12), 3482-3487.

Rothen, N., Seth, A. K., Witzel, C., & Ward, J. (2013). Diagnosing synaesthesia with online colour pickers: maximising sensitivity and specificity. *Journal of neuroscience methods*, *215*(1), 156-160.

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature neuroscience*, *10*(6), 792-797.

Rouw, R., & Scholte, H. S. (2010). Neural basis of individual differences in synesthetic experiences. *Journal of Neuroscience*, *30*(18), 6205-6213.

Rouw, R., & Scholte, H. S. (2016). Personality and cognitive profiles of a general synesthetic trait. *Neuropsychologia*, *88*, 35-48.

Rouw, R., Case, L., Gosavi, R., & Ramachandran, V. (2014). Color associations for days and letters across different languages. *Frontiers in psychology, 5*, 123-139.

Rouw, R., Scholte, H. S., & Colizoli, O. (2011). Brain areas involved in synaesthesia: a review. *Journal of neuropsychology*, *5*(2), 214-242.

Rouw, R., van Driel, J., Knip, K., & Ridderinkhof, K. R. (2013). Executive functions in synesthesia. *Consciousness and cognition, 22*(1), 184-202.

Sachs, M. C. (2017). plotROC: A Tool for Plotting ROC Curves. *Journal of Statistical Software, 79.*

Sagiv, N., & Ward, J. (2006). Crossmodal interactions: lessons from synesthesia. *Progress in brain research, 155*, 259-271.

Sagiv, N., Heer, J., & Robertson, L. (2006a). Does binding of synesthetic color to the evoking grapheme require attention? *Cortex*, *42*(2), 232-242.

Sagiv, N., Simner, J., Collins, J., Butterworth, B., & Ward, J. (2006b). What is the relationship between synaesthesia and visuo-spatial number forms?. *Cognition*, *101*(1), 114-128.

Salcicioli, J.D., Crutiain, Y., Komorwski, M., & Marshall, D.C. (2016). Sensitivity Analysis and Model Validation. In *Secondary Analysis of Electronic Health Records* (pp. 263-274). Cambridge, MA: Springer.

Salkind, N. J. (Ed.). (2010). *Encyclopedia of research design* (Vol. 1). Thousand Oaks, CA: Sage.

Sarkar, D. (2008) *Lattice: Multivariate Data Visualization with R*. New York: Springer.

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763-1768.

Schweizer, T. A., Li, Z., Fischer, C. E., Alexander, M. P., Smith, S. D., Graham, S. J., & Fornazzari, L. (2013). From the thalamus with love: a rare window into the locus of emotional synesthesia. *Neurology*, *81*(5), 509-510.

Sengül, A., Rognini, G., van Elk, M., Aspell, J. E., Bleuler, H., & Blanke, O. (2013). Force feedback facilitates multisensory integration during robotic tool use. *Experimental brain research*, *227*(4), 497-507.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature, 408*(6814), 783.

Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport, 16*(17(¡), 1923-1927.

Shore, D. I., Barnes, M. E., & Spence, C. (2006). Temporal aspects of the visuotactile congruency effect. *Neuroscience letters*, *392*(1-2), 96-100.

Shriki, O., Sadeh, Y., & Ward, J. (2016). The emergence of synaesthesia in a neuronal network model via changes in perceptual sensitivity and plasticity. *PLoS computational biology*, *12*(7), e1004959.

Simner, J. (2012). Defining synaesthesia. *British journal of psychology*, *103*(1), 1-15.

Simner, J., & Carmichael, D. A. (2015). Is synaesthesia a dominantly female trait? *Cognitive neuroscience*, *6*(2-3), 68-76.

Simner, J., & Holenstein, E. (2007). Ordinal linguistic personification as a variant of synesthesia. *Journal of cognitive neuroscience*, *19*(4), 694-703.

Simner, J., & Hubbard, E. M. (Eds.) (2013). The Oxford Handbook of Synesthesia. Oxford, UK: Oxford University Press.

Simner, J., & Logie, R. H. (2008). Synaesthetic consistency spans decades in a lexical–gustatory synaesthete. *Neurocase*, *13*(5-6), 358-365.

Simner, J., & Ludwig, V. U. (2012). The color of touch: A case of tactile-visual synaesthesia. *Neurocase, 18*(2), 167-180.

Simner, J., Carmichael, D. A., Hubbard, E. M., Morris, Z., & Lawrie, S. M. (2015). Rates of white matter hyperintensities compatible with the radiological profile of multiple sclerosis within self-referred synesthete populations. *Neurocase*, *21*(3), 322-330.

Simner, J., Gärtner, O., & Taylor, M. D. (2011). Cross-modal personality attributions in synaesthetes and non-synaesthetes. *Journal of Neuropsychology, 5*(2), 283-301.

Simner, J., Harrold, J., Creed, H., Monro, L., & Foulkes, L. (2009a). Early detection of markers for synaesthesia in childhood populations. *Brain, 132*(1), 57–64.

Simner, J., Ipser, A., Smees, R., & Álvarez, J. (2017). Does synaesthesia age? Changes in the quality and consistency of synaesthetic associations. *Neuropsychologia*, *106*, 407-416.

Simner, J., Mayo, N., & Spiller, M. J. (2009b). A foundation for savantism? Visuo-spatial synaesthetes present with cognitive benefits. *Cortex*, *45*(10), 1246-1260.

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., Scott, K., & Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, *35*(8), 1024-1033.

Simon, J. R., & Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics, 6*(1), 99-105.

Šimundić, A. M. (2009). Measures of diagnostic accuracy: basic definitions. *Ejifcc*, *9*(4), 203.

Sinke, C., Neufeld, J., Zedler, M., Emrich, H. M., Bleich, S., Münte, T. F., & Szycik, G. R. (2014). Reduced audiovisual integration in synesthesia–evidence from bimodal speech perception. *Journal of neuropsychology, 8*(1), 94-106.

Skelton, R., Ludwig, C., & Mohr, C. (2009). A novel, illustrated questionnaire to distinguish projector and associator synaesthetes. *Cortex*, *45*(6), 721-729.

Smilek, D., Callejas, A., Dixon, M. J., & Merikle, P. M. (2007a). Ovals of time: Time-space associations in synaesthesia. *Consciousness and Cognition*, *16*(2), 507-519.

Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2001). Synaesthetic photisms influence visual perception. *Journal of cognitive neuroscience*, *13*(7), 930-936.

Smilek, D., Malcolmson, K. A., Carriere, J. S. A., Eller, M., Kwan, D., & Reynolds, M. (2007b). When '3' is a jerk and 'T' is a king: Personifying inanimate objects in synaesthesia. *Journal of Cognitive Neuroscience, 19*(6), 981-992.

Specht, K. (2012). Synaesthesia: cross activations, high interconnectivity, and a parietal hub. *Translational Neuroscience, 3*(1), 15-21.

Spence, C., Kingstone, A., Shore, D. I., & Gazzaniga, M. S. (2001). Representation of visuotactile space in the split brain. *Psychological Science, 12*, 90–93.

Spence, C., Nicholls, M. E., Gillespie, N., & Driver, J. (1998). Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. *Perception & psychophysics, 60*(4), 544-557.

Spence, C., Pavani, F., & Driver, J. (2004a). Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 148-169.

Spence, C., Pavani, F., Maravita, A., & Holmes, N. (2004b). Multisensory contributions to the 3-D representation of visuotactile peripersonal space in humans: evidence from the crossmodal congruency task. *Journal of Physiology-Paris*, *98*(1-3), 171-189.

Spence, C., Pavani, F., Maravita, A., & Holmes, N. (2008). Multi-sensory interactions. In M. C. Lin & M. Otaduy (Eds.), *Haptic rendering: foundations, algorithms, and applications*. New York: AK Peters/CRC Press.

Sperling, J. M., Prvulovic, D., Linden, D. E., Singer, W., & Stirn, A. (2006). Neuronal correlates of colour-graphemic synaesthesia: AfMRI study. *Cortex*, *42*(2), 295-303.

Spiller, M. J., & Jansari, A. S. (2008). Mental imagery and synaesthesia: Is synaesthesia from internally-generated stimuli possible? *Cognition*, *109*(1), 143-151.

Spiller, M. J., Harkry, L., McCullagh, F., Thoma, V., & Jonas, C. (2019). Exploring the relationship between grapheme colour-picking consistency and mental imagery. *Philosophical Transactions of the Royal Society B, 374*(17887), 20190023.

Spiller, M. J., Jonas, C. N., Simner, J., & Jansari, A. (2015). Beyond visual imagery: How modality-specific is enhanced mental imagery in synesthesia?. *Consciousness and cognition*, *31*, 73-85.

Stekelenburg, J. J. Maes, J. P., van Gool, A. R., Sitskoorn, M., & Vroomen, J. (2013). Deficient multisensory integration in schizophrenia: an event-related potential study. *Schizophrenia research, 147*(2), 253-261.

Steven, M. S., Hansen, P. C., & Blakemore, C. (2006). Activation of color-selective areas of the visual cortex in a blind synesthete. *Cortex*, *42*(2), 304-308.

Stevens, J. R. (2017). Replicability and reproducibility in comparative psychology. *Frontiers in psychology*, *8*, 862.

Stevenson, M. (2019, August 23). *epiR: Tools for the Analysis of Epidemiological Data. R package version 0.9-96* Retrieved from https://CRAN.R-project.org/package=epiR

Stevenson, R. A., Ghose, D., Krueger Fister, J., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., Kurela, L. R., Siemann, J. K., James, T. W., & Wallace, M. T. (2014). Identifying and quantifying multisensory integration: A tutorial review. *Brain topography, 27*, 707-730.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology, 18*(6), 643.

Tabibnia, G., Monterosso, J. R., Baicy, K., Aron, A. R., Poldrack, R. A., Chakrapani, S., Lee, B., & London, E. D. (2011). Different forms of self-control share a neurocognitive substrate. *Journal of Neuroscience, 31*(13), 4805-4810.

Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., ... & Shrout, P. E. (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science*, *12*(5), 742-756.

Talsma, D., & Kok, A. (2002). Intermodal spatial attention differs between vision and audition: An event-related potential analysis. *Psychophysiology*, *39*(6), 689-706.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in cognitive sciences*, *14*(9), 400-410.

Tang, X., Wu, J., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience & Biobehavioral Reviews*, *61*, 208-224.

Teichmann, A. L., Nieuwenstein, M. R., & Rich, A. N. (2017). Digit–color synaesthesia only enhances memory for colors in a specific context: A new method of duration thresholds to measure serial recall. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(8), 1494.

Terhune, D. B., Luke, D. P., & Kadosh, R. C. (2017). The induction of synaesthesia in non-synaesthetes. *Sensory Blending: On Synaesthesia and Related Phenomena*, 215.

Terhune, D. B., Tai, S., Cowey, A., Popescu, T., & Cohen-Kadosh, R. (2011). Enhanced cortical excitability in grapheme-color synaesthesia and its modulation. *Current Biology, 21*(23), 2006-2009.

Thiele, C. (2018, September 17). *cutpointr: Determine and Evaluate Optimal Cutpoints in Binary Classification Ta sks. R package version 0.7.3*. Retrieved from https://CRAN.R-project.org/package=cutpointr

Thompson, W. L., Slotnick, S.D., Burrage, M. S., & Kosslyn, S. M. (2009). Two forms of spatial imagery: Neuroimaging evidence. *Psychological Science, 20*(10), 1245–1253.

Tiippana, K., Andersen, T. S., & Sams, M. (2004). Visual attention modulates audiovisual speech perception. *European Journal of Cognitive Psychology*, *16*(3), 457-472.

Van den Wildenberg, W. P., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., & Ridderinkhof, K. R. (2010). To head or to heed? Beyond the surface of selective action inhibition: a review. *Frontiers in human neuroscience*, *4*, 222.

Van der Stoep, N., van der Stigchel, S., & Nijboer, T. C. W. (2015). Exogenous spatial attention decreases audiovisual integration. *Attention, perception, & psychophysics*, *77*(2), 464-482.

Van der Veen, F. M., Aben, H. P., Smits, M., & Röder, C. H. (2014). Grapheme-color synesthesia interferes with color perception in a standard Stroop task. *Neuroscience, 258*, 246-253.

Van Elk, M., Forget, J., & Blanke, O. (2013). The effect of limb crossing and limb congruency on multisensory integration in peripersonal space for the upper and lower extremities. *Consciousness and cognition*, *22*(2), 545-555.

Van Leeuwen, T. M. (2013). Individual differences in synesthesia. In J. Simner & E. M. (Eds.), *The Oxford Handbook of Synesthesia* (pp. 241-264). Oxford, UK: Oxford University Press.

Van Leeuwen, T. M., Den Ouden, H. E., & Hagoort, P. (2011). Effective connectivity determines the nature of subjective experience in grapheme-color synesthesia. *Journal of Neuroscience*, *31*(27), 9879-9884.

Van Leeuwen, T. M., Petersson, K. M., & Hagoort, P. (2010). Synaesthetic colour in the brain: beyond colour areas. A functional magnetic resonance imaging study of synaesthetes and matched controls. *PLoS One, 5*(8), e12074.

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences*, *102*(4), 1181-1186.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th ed.) New York: Springer.

Verbruggen, F., Liefooghe, B., Notebaert, W., & Vandierendonck, A. (2005). Effects of stimulus–stimulus compatibility and stimulus–response compatibility on response inhibition. *Acta psychologica*, *120*(3), 307-326.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48.

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and Influence Diagnostics for Meta-Analysis. *Research Synthesis Methods, 1*(2), 112–25.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam Med*, *37*(5), 360-363.

Walton, M., & Spence, C. (1999). Cross-modal interference of touch on vision: spatial coordinates and visual capture. International Multisensory Research Forum, Oxford, UK, October 1-2.

Walton, M., & Spence, C. (2004). Cross-modal congruency and visual capture in a visual elevation-discrimination task. *Experimental Brain Research*, *154*(1), 113-120.

Ward, J. (2013). Synesthesia. *Annual review of psychology, 64*, 49-75.

Ward, J., & Banissy, M. J. (2015). Explaining mirror-touch synesthesia. *Cognitive Neuroscience, 6*(2-3), 118-133.

Ward, J., & Simner, J. (2003). Lexical-gustatory synaesthesia: linguistic and conceptual factors. *Cognition, 89*(3), 237-261.

Ward, J., Hovard, P., Jones, A., & Rothen, N. (2013). Enhanced recognition memory in grapheme-color synaesthesia for different categories of visual stimuli. *Frontiers in psychology*, *4*, 762.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex, 42*(2), 264-280.

Ward, J., Ipser, A., Phanvanova, E., Brown, P., Bunte, I., & Simner, J. (2018a). The prevalence and cognitive profile of sequence-space synaesthesia. *Consciousness and cognition*, *61*, 79-93.

Ward, J., Jonas, C., Dienes, Z., & Seth, A. (2009). Grapheme-colour synaesthesia improves detection of embedded shapes, but without pre-attentive 'pop-out'of synaesthetic colour. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1684), 1021-1026.

Ward, J., Li, R., Salih, S., & Sagiv, N. (2007). Varieties of grapheme-colour synaesthesia: a new theory of phenomenological and behavioural differences. *Consciousness and cognition*, *16*(4), 913-931.

Ward, J., Rothen, N., Chang, A., & Kanai, R. (2017a). The structure of inter-individual differences in visual ability: Evidence from the general population and synaesthesia. *Vision research*, *141*, 293-302.

Ward, J., Sagiv, N., & Butterworth, B. (2009). The impact of visuo-spatial number forms on simple arithmetic. *Cortex*, *45*(10), 1261-1265.

Ward, J., Schnakenberg, P., & Banissy, M. J. (2018b). The relationship between mirror-touch synaesthesia and empathy: New evidence and a new screening tool. *Cognitive neuropsychology*, *35*(5-6), 314-332.

Ward, J., Simner, J., & Auyeung, V. (2005). A comparison of lexical-gustatory and grapheme-colour synaesthesia. *Cognitive neuropsychology, 22*(1), 28-41.

Watson, M. R., Akins, K., Spiker, C., Crawford, L., & Enns, J. T. (2014). Synesthesia and learning: a critical review and novel theory. *Frontiers in human neuroscience*, *8*, 98.

Weiss, P. H., & Fink, G. R. (2008). Grapheme-colour synaesthetes show increased grey matter volumes of parietal and fusiform cortex. *Brain*, *132*(1), 65-70.

Weiss, P. H., Zilles, K., & Fink, G. R. (2005). When visual perception causes feeling: enhanced cross-modal processing in grapheme-color synesthesia. *Neuroimage*, *28*(4), 859-868.

Whittingham, K. M., McDonald, J. S., & Clifford, C. W. (2014). Synesthetes show normal sound-induced flash fission and fusion illusions. *Vision research, 105*, 1-9.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

Witthoft, N., & Winawer, J. (2006). Synesthetic colors determined by having colored refrigerator magnets in childhood. *Cortex*, *42*(2), 175-183.

Witthoft, N., & Winawer, J. (2013). Learning, memory, and synesthesia. *Psychological science*, *24*(3), 258-265.

Yaro, C., & Ward, J. (2007). Searching for Shereshevskii: What is superior about the memory of synaesthetes? *The Quarterly Journal of Experimental Psychology*, *60*(5), 681-695.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32-35.

Zopf, R., Savage, G., & Williams, M. A. (2010). Crossmodal congruency measures of lateral distance effects on the rubber hand illusion. *Neuropsychologia*, *48*(3), 713-725.

Zopf, R., Savage, G., & Williams, M. A. (2013). The crossmodal congruency task as a means to obtain an objective behavioral measure in the rubber hand illusion paradigm. *JoVE (Journal of Visualized Experiments), 77*, e50530.

Appendices

Appendix A: Edinburgh Synaesthesia Screening Assessment (ESSA Pilot Version)

Edinburgh Synaesthesia Screening Assessment (ESSA)

pilot version – interview administration

Subject ID	
Date	

Date of birth (DD/MM/YY)	
Current residence (City, Country)	
Usual residence (City, Country) *If different*	
Nationality	
Gender	Female Male Other (specify) Rather not say
Handedness	Right-handed Left-handed Ambidextrous
Which is/are your native language/s? (I.e. Language/s you are proficient at and to which you were exposed before starting school)	
Are you proficient at any other language/s?	
Do you consider yourself multilingual? (I.e. You think habitually, or in certain contexts, in more than one language)	Yes No Not sure Rather not say
Highest level of education achieved or current level of education	
Main specialisation of your studies	
Occupation (can be student)	

- Start each block reading the descriptive statement-question to the interviewee and then proceed to read aloud each different trigger individually.
- The interviewee's direct answer corresponds to column **A** (Automaticity) and should be coded as follows: *Always* (2), *Sometimes* (1), *Never* (0).
- If the answer to **A** is *Never*, complete the rest of the columns with 0s.
- If the answer to **A** is *Always* or *Sometimes*, further ask the individual the following questions:
  - **C** (Constancy) – Do you always experience the same association? E.g. Is letter A always coloured red or does the colour change? *Consistent* (2), *Variable* (1)
  - **L** (Location) – Do you experience it outside your body or in your mind's eye? *Projector* (2), *Associator* (1)
  - **S** (Stability) – Have you always had this experience? *Always that I can remember* (2), *It started after some specific event/point in time* (1)

1) I always/sometimes/never see different COLOURS when I see/hear/feel/think about…	A	C	L	S
1.1) Letters				
1.2) Words				
1.3) Numbers				
1.4) Weekdays				
1.5) Months				
1.6) Visual patterns or shapes				
1.7) People's personalities				
1.8) Emotions				
1.9) Orgasms				
1.10) Touch				
1.11) Pain				
1.12) Temperatures				
1.13) Tastes				
1.14) Smells				
1.15) Sounds				
1.16) Music				
1.17) Movements				
1.18) Other (specify)				

2) I always/sometimes/never see different VISUAL PATTERNS when I see/hear/feel/think about…	A	C	L	S
2.1) Colours				
2.2) People's voices				
2.3) Emotions				
2.4) Touch				
2.5) Pain				
2.6) Temperatures				
2.7) Tastes				
2.8) Smells				
2.9) Sounds				
2.10) Music				
2.11) Other (specify)				

3) TICKER-TAPE	A	C	L	S
3.1) I always/sometimes/never visualise spoken words or thoughts as if they were subtitles.				

4) I always/sometimes/never see the following items as being arranged in SPECIFIC PATTERNS IN SPACE:	A	C	L	S
4.1) Alphabet				
4.2) Weekdays				
4.3) Months				
4.4) Numbers				
4.5) Musical notes				
4.6) Other (specify)				

5) I always/sometimes/never experience an actual physical sensation of TOUCH or PAIN when I see/hear/feel/think about…	A	C	L	S
5.1) Letters				
5.2) Words				
5.3) Numbers				
5.4) Colours				
5.5) Visual patterns or shapes				
5.6) Emotions				
5.7) Tastes				
5.8) Smells				
5.9) Sounds				
5.10) Music				
5.11) Other (specify)				

6) I always/sometimes/never experience physical TEMPERATURE changes when I see/hear/feel/think about…	A	C	L	S
6.1) Colours				
6.2) Visual patterns or shapes				
6.3) Emotions				
6.4) Touch				
6.5) Pain				
6.6) Tastes				
6.7) Smells				
6.8) Sound				
6.9) Music				
6.10) Other (specify)				

7) I always/sometimes/never experience an actual physical sensation of TASTE in my mouth when I see/hear/feel/think about…	A	C	L	S
7.1) Colours				
7.2) Visual patterns or shapes				
7.3) Letters				
7.4) Words				
7.5) Numbers				
7.6) People's personalities				
7.7) Emotions				
7.8) Orgasms				
7.9) Touch				
7.10) Pain				
7.11) Sounds				
7.12) Music				
7.13) Other (specify)				

8) I always/sometimes/never experience an actual physical sensation of SMELL in my nose when I see/hear/feel/think about…	A	C	L	S
8.1) Colours				
8.2) Visual patterns or shapes				
8.3) Letters				
8.4) Words				
8.5) Numbers				
8.6) People's personalities				
8.7) Emotions				
8.8) Orgasms				
8.9) Touch				
8.10) Pain				
8.11) Sounds				
8.12) Music				
8.13) Other (specify)				

9) MIRRORS	A	C	L	S
9.1) When I see/think about someone else or something else being touched, I always/sometimes/never experience the same actual physical sensation of touch on myself.				
9.2) When I see/think about someone else or something else being hurt, I always/sometimes/never experience the same actual physical sensation of pain on myself.				
9.3) When I see/think about someone else tasting something, I always/sometimes/never experience the same actual physical sensation of taste in my mouth.				
9.4) When I see/think about someone else smelling something, I always/sometimes/never experience the same actual physical sensation of smell in my nose.				
9.5) When I see/think about someone else feeling a specific emotion, I always/sometimes/never experience the same actual emotion myself.				
9.6) Other (specify)				

10) I always/sometimes/never hear actual SOUNDS or MUSIC when I see/hear/feel/think about…	A	C	L	S
10.1) Letters				
10.2) Words				
10.3) Weekdays				
10.4) Months				
10.5) Numbers				
10.6) Emotions				
10.7) Touch				
10.8) Pain				
10.9) Temperatures				
10.10) Tastes				
10.11) Smells				
10.12) Movements				
10.13) Other (specify)				

11) To me, the following items have established PERSONALITIES and/or GENDERS:	A	C	L	S
11.1) Letters				
11.2) Numbers				
11.3) Movements				
11.4) Objects				
11.5) Other (specify)				

Appendix B: Studies 1, 2, and 3 Extended Bayesian Interference Analyses

Study 1

The error rates (ER) analyses revealed that the 'Congruency'-only models were the best performing ones and extremely preferred to the null model for both tasks (Cross-modal Congruency Task or CCT: BF_{10} = 7760.06; Flanker Task or FT: BF_{10} = 2.598.44). There was also decisive or extreme evidence for the alternative hypothesis for the non-interaction and interaction models (CCT: BF_{10} = 2381.34 and BF_{10} = 886.294, respectively; FT: BF_{10} = 1654.53 and BF_{10} = 1162.53, respectively). However, there was moderate or substantial evidence for the exclusion of the factor 'Group' in the CCT (Inclusion BF_{10} = .307), but the inclusion/exclusion of the interaction term presented inconclusive evidence (Inclusion BF_{10} = .372). There was inconclusive evidence for both the factor 'Group' and the interaction term in the FT (Inclusion BF_{10} = .637 and Inclusion BF_{10} = .703, respectively).

Regarding the reaction times (RT) analyses, the interaction model was the best performing and extremely preferred to the null model for the CCT (BF_{10} = 6.552e+10) and there was extreme evidence as well in favour of the alternative hypothesis for the 'Congruency'-only and non-interaction models (BF_{10} = 1.111e+10 and BF_{10} = 2.236e+10, respectively). The 'Group'-only model presented inconclusive evidence (BF_{10} = 1.934). On the other hand, in the FT the 'Congruency'-only model was the best performing and the one preferred to the null model with extreme evidence (BF_{10} = 2.673e+16). The non-interaction and the interaction models also presented extreme evidence in favour of the alternative hypothesis (BF_{10} = 2.593e+16 and BF_{10} = 1.151e+16, respectively), and there was inconclusive evidence for the 'Group'-only model (BF_{10} = .658). However, in both tasks, the inclusion/exclusion of the factor 'Group' and the interaction term presented inconclusive or anecdotal evidence (CCT: Inclusion BF_{10} = 2.012 and Inclusion BF_{10} = 2.931, respectively; FT: Inclusion BF_{10} = .97 and Inclusion BF_{10} = .444, respectively).

Study 2

The ER analyses for the visual Unimodal Congruency Task (vUCT) task provided anecdotal or inconclusive evidence for the ‘Congruency’-only and ‘Group’-only models (BF_{10} = .443 and BF_{10} = .446, respectively) and evidence in favour of the null hypothesis for the models involving the two factors: substantial evidence for the non-interaction model (BF_{10} = .197) and strong for the interaction model (BF_{10} = .07). However, there was anecdotal evidence for the inclusion/exclusion of the factor ‘Group’ and the interaction term (Inclusion BF_{10} = .445 and Inclusion BF_{10} = .356). Similarly, in the reversed Cross-modal Congruency Task (rCCT) task, the interaction model presented substantial evidence in favour of the null hypothesis (BF_{10} = .186) and there was anecdotal or inconclusive evidence for the rest of the models (‘Congruency’-only: BF_{10} = .740; ‘Group’-only: BF_{10} = .821; non-interaction: BF_{10} = .564). Inconclusive evidence was also observed for the inclusion/exclusion of the factor ‘Group’ and the interaction term (Inclusion BF_{10} = .796 and Inclusion BF_{10} = .329, respectively).

The RT analyses revealed that the ‘Congruency’-only models were the best performing ones and preferred to the null model with moderate or substantial evidence for the vUCT task (BF_{10} = 8.86) and extreme or decisive evidence for the rCCT (BF_{10} = 2852.57). There was also moderate evidence for the alternative hypothesis for the non-interaction and interaction models in the vUCT (BF_{10} = 7.909 and BF_{10} = 3.208, respectively) and extreme evidence in the rCCT (BF_{10} = 2352.2 and BF_{10} = 768.013, respectively). The ‘Group’-only model presented inconclusive evidence in both tasks (vUCT: BF_{10} = .818; rCCT: BF_{10} = .816). The inclusion/exclusion of the factor ‘Group’ and the interaction term also presented inconclusive evidence (vUCT: Inclusion BF_{10} = .885 and Inclusion BF_{10} = .406, respectively; rCCT: Inclusion BF_{10} = .825 and BF_{10} = .327, respectively).

Study 3

The ER analyses for the VTCU task provided anecdotal or inconclusive evidence for the ‘Group’-only model (BF_{10} = .557) and moderate to strong evidence in favour of the null hypothesis for the other models ('Congruency'-only: BF_{10} = .223; non-interaction: BF_{10} = .122; interaction: BF_{10} = .036). Accordingly, there was anecdotal evidence for the inclusion/exclusion of the factor 'Group' (Inclusion BF_{10} = .555) and moderate evidence against the inclusion of the interaction term (Inclusion BF_{10} = .295). In the AVCU task, the analyses provided anecdotal evidence for the 'Congruency'-only model (BF_{10} = .584) and moderate to strong evidence in favour of the null hypothesis for the other models ('Group'-only: BF_{10} = .233; non-interaction: BF_{10} = .138; interaction: BF_{10} = .055). Similarly, there was moderate evidence against the inclusion of the factor 'Group' (Inclusion BF_{10} = .234) and anecdotal evidence for the interaction term (Inclusion BF_{10} = .395). Regarding the VTCR and AVCR tasks, the analyses determined that the 'Congruency'-only models were the best performing ones and strongly preferred to the null models (VTCR: BF_{10} = 1.176e+9 and AVCR: BF_{10} = 7.680e+8). There was also moderate evidence in favour of the null hypothesis for the 'Group'-only models (VTCR: BF_{10} = .112 and AVCR: BF_{10} = .130) and extreme evidence in favour of the alternative hypothesis for the non-interaction (VTCR: BF_{10} = 1.494e+8 and AVCR: BF_{10} = 1.296e+8) and the interaction models (VTCR: BF_{10} = 2.086e+7 and AVCR: BF_{10} = 2.909e+7). However, the exclusion of the factors 'Group' and the interaction terms was supported by substantial evidence (VTCR: Inclusion BF_{10} = .127 and Inclusion BF_{10} = .140, respectively; AVCR: Inclusion BF_{10} = .169 and Inclusion BF_{10} = .224, respectively).

The RT analyses revealed that the 'Congruency'-only models were the best performing ones and strongly preferred to the null models for all tasks (VTCU: BF_{10} = 1.282e+8; AVCU: BF_{10} = 2.023e+11; VTCR: BF_{10} = 1.797e+19; AVCR: BF_{10} = 9.550e+15). There was also moderate evidence in favour of the null hypothesis for the 'Group'-only models for the VTCR and AVCR tasks (BF_{10} = .303 and BF_{10} = .269, respectively), but the evidence was inconclusive for the VTCU and AVCU tasks (BF_{10} = .546 and BF_{10} = .642). The non-interaction and interaction models presented extreme evidence in favour of the alternative hypothesis for all tasks (VTCU: BF_{10} = 8.592e+7 and BF_{10} = 1.626e+8, respectively; AVCU: BF_{10} = 1.117e+11 and BF_{10} = 1.741e+10, respectively; VTCR: BF_{10} = 1.006e+19 and BF_{10} = 1.616e+18, respectively; AVCR: BF_{10} = 4.645e+15 and BF_{10} = 1.926e+15, respectively). However, there was substantial evidence supporting the exclusion of the interaction term in the AVCU and VTCR tasks (Inclusion BF_{10} = .156 and Inclusion BF_{10} = .161, respectively) and the evidence was inconclusive for the VTCU and AVCR tasks (Inclusion BF_{10} = 1.893 and Inclusion BF_{10} = .415). Similarly, the evidence for the inclusion/exclusion of the factor ‘Group’ was inconclusive for all tasks (VTCU: Inclusion BF_{10} = .670; AVCU: Inclusion BF_{10} = .552; VTCR: Inclusion BF_{10} = .560; AVCR: Inclusion BF_{10} = .486).

Appendix C: Studies 1, 2, and 3 Multivariate Mixed Model Additional Tables

For each study and task, we first compared the target or full model (i.e. accuracy rates and reactiones times as a function of the interaction between ‘Congruency’ and ‘Group’) to the null model (i.e. accuracy rates/reaction times-only models), the ‘Congruency’-only model (i.e. model without the ‘Group’ term), and the no-random effects model (i.e. full model without random effects) (Tables C1, C3, and C5). Following next, we proceeded to analyse the components of the full model (Tables C2, C4, and C6).

Table C1.
Multivariate mixed model analyses: model comparisons statistics (Study 1).

Model comparisons	Statistics
Accuracy rates analyses	
Full model (F) vs. Null model (N)	
CCT	AIC = 2068 (F), AIC = 2111 (N); $\chi2(3) = 48.8$, $p < .001$
FT	AIC = 896 (F), AIC = 938 (N); $\chi2(3) = 26$, $p < .001$
Full model (F) vs. 'Congruency'-only model (C)	
CCT	AIC = 2068 (F), AIC = 2066 (C); $\chi2(2) = 2.67$, $p = .26$
FT	AIC = 896 (F), AIC = 894 (C); $\chi2(2) = 2.18$, $p = .34$
Full model (F) vs. No-random effects model (R)	
CCT	AIC = 2068 (F), AIC = 2408 (R); $\chi2(3) = 346$, $p < .001$
FT	AIC = 896 (F), AIC = 966 (R); $\chi2(3) = 76$, $p < .001$
Reaction times analyses	
Full model (F) vs. Null model (N)	
CCT	AIC = 81004 (F), AIC = 81067 (N); $\chi2(3) = 68$, $p < .001$
FT	AIC = 35492 (F), AIC = 35576 (N); $\chi2(3) = 90$, $p < .001$
Full model (F) vs. 'Congruency'-only model (C)	
CCT	AIC = 81004 (F), AIC = 81008 (C); $\chi2(2) = 7.29$, $p = .026$
FT	AIC = 35492 (F), AIC = 34492 (C); $\chi2(2) = 3.38$, $p = .18$
Full model (F) vs. No-random effects model (R)	
CCT	AIC = 81004 (F), AIC = 84115 (R); $\chi2(3) = 3117$, $p < .001$
FT	AIC = 35492 (F), AIC = 36342 (R); $\chi2(3) = 855$, $p < .001$

Note: CCT = Cross-modal Congruency Task; FT = Flanker Task; AIC = Akaike Information Criteria.

Table C2.
Multivariate mixed model analyses: full model statistics (Study 1).

Model specifications	Statistics
Accuracy rates analyses	
CCT	
Random effects	
Participant - Intercept (Congruent)	σ2 = 1.59, SD = 1.26
Participant - Incongruent	σ2 = .89, SD = .95, r = -.36
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 6.20, SE = .66, z = 9.33, p < .001
Incongruent	β = -3.40, SE = .64, z = -5.28, p < .001
Synaesthetes	β = -.97, SE = .69, z = -1.40, p = .16
Incongruent:Synaesthetes	β = .98, SE = .65, z = 1.51, p = .13
FT	
Random effects	
Participant - Intercept (Congruent)	σ2 = .17, SD = .42
Participant - Incongruent	σ2 = .60, SD = .78, r = 1
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 6.14, SE = .72, z = 8.58, p < .001
Incongruent	β = -2.76, SE = .76, z = -3.64, p < .001
Synaesthetes	β = -.96, SE = .84, z = -1.15, p = .26
Incongruent:Synaesthetes	β = .41, SE = .89, z = .46, p = .65
Reaction times analyses	
CCT	
Random effects	
Participant - Intercept (Congruent)	σ2 = 12961, SD = 114
Participant - Incongruent	σ2 = 3649, SD = 60.4, r = .32
Residual	σ2 = 22778, SD = 151
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 580, SE = 27.1, t(35) = 21.4, p < .001
Incongruent	β = 185, SE = 15.2, t(35.1) = 12.2, p < .001
Synaesthetes	β = -38.9, SE = 38.9, t(35) = -1, p = .32
Incongruent:Synaesthetes	β = -61.9, SE = 21.8, t(35.2) = -2.84, p = .008
FT	
Random effects	
Participant - Intercept (Congruent)	σ2 = 1136, SD = 33.7
Participant - Incongruent	σ2 = 237, SD = 15.4, r = .03
Residual	σ2 = 3435, SD = 58.6
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 328, SE = 8.19, t(35) = 40, p < .001
Incongruent	β = 63.4, SE = 4.63, t(35) = 13.7, p < .001
Synaesthetes	β = 21.5, SE = 11.8, t(35) = 1.83, p = .08
Incongruent:Synaesthetes	β = 1.96, SE = 6.66, t(35.4) = .29, p = .77

Note: CCT = Cross-modal Congruency Task; FT = Flanker Task.

Table C3.
Multivariate mixed model analyses: model comparisons statistics (Study 2).

Model comparisons	Statistics
Accuracy rates analyses	
Full model (F) vs. Null model (N)	
vUCT	AIC = 1244 (F), AIC = 1241 (N); $\chi2(3) = 3.94$, $p = .27$
rCCT	AIC = 817 (F), AIC = 816 (N); $\chi2(3) = 5.12$, $p = .16$
Full model (F) vs. ‘Congruency’-only model (C)	
vUCT	AIC = 1244 (F), AIC = 1241 (C); $\chi2(2) = 1.60$, $p = .45$
rCCT	AIC = 817 (F), AIC = 851 (C); $\chi2(2) = 1.19$, $p = .55$
Full model (F) vs. No-random effects model (R)	
vUCT	AIC = 1244 (F), AIC = 1355 (R); $\chi2(3) = 118$, $p < .001$
rCCT	AIC = 817 (F), AIC = 876 (R); $\chi2(3) = 65.2$, $p < .001$
Reaction times analyses	
Full model (F) vs. Null model (N)	
vUCT	AIC = 155128 (F), AIC = 155133 (N); $\chi2(3) = 11.4$, $p = .009$
rCCT	AIC = 153763 (F), AIC = 153778 (N); $\chi2(3) = 21.1$, $p < .001$
Full model (F) vs. ‘Congruency’-only model (C)	
vUCT	AIC = 155128 (F), AIC = 155124 (C); $\chi2(2) = .62$, $p = .73$
rCCT	AIC = 153763 (F), AIC = 153760 (C); $\chi2(2) = .56$, $p = .75$
Full model (F) vs. No-random effects model (R)	
vUCT	AIC = 155128 (F), AIC = 160103 (R); $\chi2(3) = 4981$, $p < .001$
rCCT	AIC = 153763 (F), AIC = 159566 (R); $\chi2(3) = 5809$, $p < .001$

Note: vUCT = visual Unimodal Congruency Task; rCCT = reversed Cross-modal Congruency Task; AIC = Akaike Information Criteria.

Table C4.
Multivariate mixed model analyses: full model statistics (Study 2).

Model specifications	Statistics
Accuracy rates analyses	
vUCT	
Random effects	
Participant - Intercept (Congruent)	σ2 = 2.55, SD = 1.60
Participant - Incongruent	σ2 = .79, SD = .89, r = -.66
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 5.61, SE = .49, *z* = 11.4, *p* < .001
Incongruent	β = -.36, SE = .47, *z* = -.78, *p* = .45
Synaesthetes	β = .68, SE = .69, *z* = 1, *p* = .32
Incongruent:Synaesthetes	β = -.70, SE = .60, *z* = -1.17, *p* = .24
rCCT	
Random effects	
Participant - Intercept (Congruent)	σ2 = 2.06, SD = 1.44
Participant - Incongruent	σ2 = .10, SD = .31, r = -1
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 6.90, SE = .60, *z* = 11.5, *p* < .001
Incongruent	β = -1.04, SE = .55, *z* = -1.91, *p* = .056
Synaesthetes	β = -.77, SE = .69, *z* = -1.11, *p* = .27
Incongruent:Synaesthetes	β = .40, SE = .58, *z* = .69, *p* = .49
Reaction times analyses	
vUCT	
Random effects	
Participant - Intercept (Congruent)	σ2 = 5664, SD = 75.3
Participant - Incongruent	σ2 = 15.8, SD = 3.98, r = 1
Residual	σ2 = 11832, SD = 109
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 511, SE = 16.9, *t*(41) = 30.2, *p* < .001
Incongruent	β = 8.43, SE = 2.92, *t*(229) = 2.89, *p* = .004
Synaesthetes	β = -8.52, SE = 23.6, *t*(41) = -.36, *p* = .72
Incongruent:Synaesthetes	β = -3.11, SE = 4.06, *t*(226) = -.77, *p* = .44
rCCT	
Random effects	
Participant - Intercept (Congruent)	σ2 = 5286, SD = 72.7
Participant - Incongruent	σ2 = 40.3, SD = 6.35, r = .06
Residual	σ2 = 8855, SD = 94.1
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 475, SE = 16.34, *t*(41) = 29.1, *p* < .001
Incongruent	β = 11, SE = 2.76, *t*(41.1) = 4, *p* < .001
Synaesthetes	β = -11.9, SE = 22.8, *t*(41) = -.52, *p* = .60
Incongruent:Synaesthetes	β = -2.04, SE = 3.87, *t*(41.5) = -.53, *p* = .60

Note: vUCT = visual Unimodal Congruency Task; rCCT = reversed Cross-modal Congruency Task.

Table C5.
Multivariate mixed model analyses: model comparisons statistics (Study 3).

Model comparisons	Statistics
Accuracy rates analyses	
Full model (F) vs. Null model (N)	
VTCU	AIC = 1051 (F), AIC = 1047 (N); $\chi 2(5) = 5.91$, $p = .32$
AVCU	AIC = 782 (F), AIC = 776 (N); $\chi 2(5) = 4.26$, $p = .51$
VTCR	AIC = 3795 (F), AIC = 3864 (N); $\chi 2(5) = 79.2$, $p < .001$
AVCR	AIC = 4138 (F), AIC = 4225 (N); $\chi 2(5) = 96.6$, $p < .001$
Full model (F) vs. 'Congruency'-only model (C)	
VTCU	AIC = 1051 (F), AIC = 1048 (C); $\chi 2(4) = 4.70$, $p = .32$
AVCU	AIC = 782 (F), AIC = 778 (C); $\chi 2(4) = 4.07$, $p = .40$
VTCR	AIC = 3795 (F), AIC = 3790 (C); $\chi 2(4) = 2.78$, $p = .60$
AVCR	AIC = 4138 (F), AIC = 4132 (C); $\chi 2(4) = 1.58$, $p = .81$
Full model (F) vs. No-random effects model (R)	
VTCU	AIC = 1051 (F), AIC = 1068 (R); $\chi 2(3) = 22.8$, $p < .001$
AVCU	AIC = 782 (F), AIC = 874 (R); $\chi 2(3) = 97.8$, $p < .001$
VTCR	AIC = 3795 (F), AIC = 4163 (R); $\chi 2(3) = 330$, $p < .001$
AVCR	AIC = 4138 (F), AIC = 5354 (R); $\chi 2(3) = 1221$, $p < .001$
Reaction times analyses	
Full model (F) vs. Null model (N)	
VTCU	AIC = 135885 (F), AIC = 135925 (N); $\chi 2(5) = 50.4$, $p < .001$
AVCU	AIC = 130042 (F), AIC = 130139 (N); $\chi 2(5) = 70.6$, $p < .001$
VTCR	AIC = 136129 (F), AIC = 136230 (N); $\chi 2(5) = 111$, $p < .001$
AVCR	AIC = 131089 (F), AIC = 131196 (N); $\chi 2(5) = 80.6$, $p < .001$
Full model (F) vs. 'Congruency'-only model (C)	
VTCU	AIC = 135885 (F), AIC = 135889 (C); $\chi 2(4) = 12$, $p = .017$
AVCU	AIC = 130042 (F), AIC = 130037 (C); $\chi 2(4) = 3.08$, $p = .54$
VTCR	AIC = 136129 (F), AIC = 136123 (C); $\chi 2(4) = 1.75$, $p = .78$
AVCR	AIC = 131089 (F), AIC = 131085 (C); $\chi 2(4) = 4.05$, $p = .40$
Full model (F) vs. No-random effects model (R)	
VTCU	AIC = 135885 (F), AIC = 142140 (R); $\chi 2(3) = 6261$, $p < .001$
AVCU	AIC = 130042 (F), AIC = 136690 (R); $\chi 2(3) = 6654$, $p < .001$
VTCR	AIC = 136129 (F), AIC = 141128 (R); $\chi 2(3) = 5005$, $p < .001$
AVCR	AIC = 131089 (F), AIC = 137322 (R); $\chi 2(3) = 6238$, $p < .001$

Note: VTCU = Visuo-Tactile Concurrent-Unrelated; AVCU = Audio-Visual Concurrent-Unrelated; VTCR = Visuo-Tactile Concurrent-Related; AVCR = Audio-Visual Concurrent-Related; AIC = Akaike Information Criteria.

Table C6.
Multivariate mixed model analyses: full model statistics (Study 3).

Model specifications	Statistics
Accuracy rates analyses	
VTCU	
Random effects	
Participant - Intercept (Congruent)	σ2 = 1, SD = 1
Participant - Incongruent	σ2 = .11, SD = .34, r = -.39
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 5.45, SE = .41, *z* = 13.3, *p* < .001
Incongruent	β = -.40, SE = .46, *z* = -.87, *p* = .39
Colour-synaesthetes	β = .65, SE = .59, *z* = 1.12, *p* = .26
Sequence-synaesthetes	β = -.58, SE = .48, *z* = -1.21, *p* = .23
Incongruent:Colour-synaesthetes	β = -.37, SE = .62, *z* = -.60, *p* = .55
Incongruent:Sequence-synaesthetes	β = .44, SE = .49, *z* = .90, *p* = .37
AVCU	
Random effects	
Participant - Intercept (Congruent)	σ2 = .56, SD = .75
Participant - Incongruent	σ2 = 2.06, SD = 1.44, r = -.19
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 5.99, SE = .70, *z* = 8.53, *p* < .001
Incongruent	β = .31, SE = .89, *z* = .35, *p* = .73
Colour-synaesthetes	β = .61, SE = .73, *z* = .83, *p* = .41
Sequence-synaesthetes	β = -.10, SE = .61, *z* = -.16, *p* = .87
Incongruent:Colour-synaesthetes	β = -1.21, SE = .96, *z* = -1.26, *p* = .21
Incongruent:Sequence-synaesthetes	β = -1.05, SE = .85, *z* = -1.23, *p* = .22
VTCR	
Random effects	
Participant - Intercept (Congruent)	σ2 = 1.27, SD = 1.13
Participant - Incongruent	σ2 = .72, SD = .85, r = -.39
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 5.22, SE = .39, *z* = 13.4, *p* < .001
Incongruent	β = -2.37, SE = .38, *z* = -6.32, *p* < .001
Colour-synaesthetes	β = .01, SE = .50, *z* = .03, *p* = .98
Sequence-synaesthetes	β = -.53, SE = .48, *z* = -1.10, *p* = .27
Incongruent:Colour-synaesthetes	β = .05, SE = .47, *z* = .11, *p* = .92
Incongruent:Sequence-synaesthetes	β = .06, SE = .45, *z* = .15, *p* = .86
AVCR	
Random effects	
Participant - Intercept (Congruent)	σ2 = 4.68, SD = 2.16
Participant - Incongruent	σ2 = 2.25, SD = 1.50, r = -.67
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	β = 5.80, SE = .61, *z* = 9.58, *p* < .001
Incongruent	β = -3.34, SE = .53, *z* = -6.27, *p* < .001
Colour-synaesthetes	β = .55, SE = .81, *z* = .66, *p* = .50
Sequence-synaesthetes	β = .32, SE = .81, *z* = -.40, *p* = .69
Incongruent:Colour-synaesthetes	β = -.22, SE = .68, *z* = -.32, *p* = .75
Incongruent:Sequence-synaesthetes	β = .51, SE = .68, *z* = -.75, *p* = .45

Continued

Reaction times analyses	
VTCU	
Random effects	
Participant - Intercept (Congruent)	σ2 = 8948, SD = 94.6
Participant - Incongruent	σ2 = 136, SD = 11.7, r = .31
Residual	σ2 = 11477, SD = 107
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	$\beta = 465$, SE = 18, $t(71) = 25.8$, $p < .001$
Incongruent	$\beta = 28$, SE = 3.92, $t(72) = 7.13$, $p < .001$
Colour-synaesthetes	$\beta = 7.50$, SE = 27.2, $t(71) = .28$, $p = .78$
Sequence-synaesthetes	$\beta = -19.7$, SE = 27.5, $t(71) = -.72$, $p = .48$
Incongruent:Colour-synaesthetes	$\beta = -9$, SE = 5.90, $t(72) = -1.53$, $p = .13$
Incongruent:Sequence-synaesthetes	$\beta = -21.1$, SE = 5.98, $t(71) = -3.53$, $p < .001$
AVCU	
Random effects	
Participant - Intercept (Congruent)	σ2 = 6790, SD = 82.4
Participant - Incongruent	σ2 = 29.6, SD = 5.44, r = -.13
Residual	σ2 = 7607, SD = 87.2
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	$\beta = 431$, SE = 15.7, $t(71.1) = 27.5$, $p < .001$
Incongruent	$\beta = 15.9$, SE = 2.86, $t(68.9) = 5.55$, $p < .001$
Colour-synaesthetes	$\beta = -13.2$, SE = 23.6, $t(71.1) = -.56$, $p = .58$
Sequence-synaesthetes	$\beta = -25.3$, SE = 24, $t(71) = -1.06$, $p = .29$
Incongruent:Colour-synaesthetes	$\beta = 5.87$, SE = 4.30, $t(68.7) = 1.37$, $p = .18$
Incongruent:Sequence-synaesthetes	$\beta = 5.19$, SE = 4.32, $t(66) = 1.20$, $p = .24$
VTCR	
Random effects	
Participant - Intercept (Congruent)	σ2 = 12013, SD = 110
Participant – Incongruent	σ2 = 3632, SD = 60.3, r = .24
Residual	σ2 = 22329, SD = 149
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	$\beta = 571$, SE = 21, $t(71.1) = 27.2$, $p < .001$
Incongruent	$\beta = 128$, SE = 12.3, $t(68.5) = 10.4$, $p < .001$
Colour-synaesthetes	$\beta = 19.7$, SE = 31.6, $t(71) = .62$, $p = .54$
Sequence-synaesthetes	$\beta = -23.8$, SE = 32, $t(71) = -.74$, $p = .46$
Incongruent:Colour-synaesthetes	$\beta = -.37$, SE = 18.5, $t(68.2) = -.02$, $p = .98$
Incongruent:Sequence-synaesthetes	$\beta = -8.65$, SE = 18.8, $t(68.4) = -.46$, $p = .65$
AVCR	
Random effects	
Participant - Intercept (Congruent)	σ2 = 17856, SD = 134
Participant – Incongruent	σ2 = 5231, SD = 72.3, r = .26
Residual	σ2 = 21181, SD = 146
Fixed effects	
Intercept (Congruent, Non-synaesthetes)	$\beta = 579$, SE = 25.5, $t(71) = 22.8$, $p < .001$
Incongruent	$\beta = 97.6$, SE = 14.6, $t(65.2) = 6.68$, $p < .001$
Colour-synaesthetes	$\beta = 17.1$, SE = 38.4, $t(71) = .44$, $p = .66$
Sequence-synaesthetes	$\beta = -35.9$, SE = 38.9, $t(71) = -.92$, $p = .36$
Incongruent:Colour-synaesthetes	$\beta = 11.4$, SE = 21.9, $t(64.5) = .52$, $p = .61$
Incongruent:Sequence-synaesthetes	$\beta = 28.8$, SE = 22.2, $t(64.7) = 1.30$, $p = .20$

Note: VTCU = Visuo-Tactile Concurrent-Unrelated; AVCU = Audio-Visual Concurrent-Unrelated; VTCR = Visuo-Tactile Concurrent-Related; AVCR = Audio-Visual Concurrent-Related.

Appendix D: Edinburgh Synaesthesia Screening Assessment (ESSA Revised Version)

Edinburgh Synaesthesia Screening Assessment (ESSA)

revised version – questionnaire administration

Subject ID	
Date	

Date of birth (DD/MM/YY)	
Current residence (City, Country)	
Usual residence (City, Country) *If different*	
Nationality	
Gender	Female Male Other (specify) Rather not say
Handedness	Right-handed Left-handed Ambidextrous
Which is/are your native language/s? (I.e. Language/s you are proficient at and to which you were exposed before starting school)	
Are you proficient at any other language/s?	
Do you consider yourself multilingual? (I.e. You think habitually, or in certain contexts, in more than one language)	Yes No Not sure Rather not say
Highest level of education achieved or current level of education	
Main specialisation of your studies	
Occupation (can be student)	

You will now read a series of statements about particular perceptual experiences. These experiences are characterised by the fact that a sensation in one of the senses, such as hearing, triggers an automatic and involuntary sensation in another sense, such as vision. For example, hearing music causes the perception of different colours.

The statements REFER to associations that are automatic and involuntary. They can be experienced either outside your body, on your body, or in your mind's eye. For instance, if you see colours for letters and you read a book, you can either see the different colours superimposed on the printed letters or 'see' them in your mind.

The statements do NOT REFER to mere cognitive or cultural associations (e.g. "Hot is red and cold is blue"), associations learned in school or other contexts (e.g. "Wednesdays are blue" because when I was a child, I learned the weekdays with a poster in which Wednesdays were coloured blue), or to sensations induced by drugs or artificial devices.

Keeping all this in mind, **carefully read each statement and decide how much they apply to you by rating them on a 5-point scale ranging from *Not at all* to *Completely***. Give your answers thinking about experiences that you HAVE, not thinking whether you could come up with an answer if needed. If a statement sounds weird or does not make sense to you, the answer is most likely *Not at all*.

For example, you will read something like "I perceive different colours when I see, hear, or think about the numbers" and you will have to decide then how much this statement applies to you by circling the appropriate rating value:

Not at all A little bit Moderately Quite a lot Completely

1.A11) I perceive different <u>colours</u> when I see, hear, or think about concepts such as letters, words, names, shapes, symbols, numbers, weekdays, months, hours, years, or measurement scales. For example: *When I think about the letter A, I perceive the colour red*, or *When I hear the word "Tuesday", I perceive a canary yellow colour*.

Not at all   A little bit   Moderately   Quite a lot   Completely

If you have answered *Not at all*, skip to Q16.

2.A11-1) I perceive different <u>colours</u> when I see, hear, or think about the letters of the alphabet: **Latin alphabet** (i.e. alphabet used by languages such as English, French, German, etc.).

Not at all   A little bit   Moderately   Quite a lot   Completely

3.A11-2) I perceive different <u>colours</u> when I see, hear, or think about the letters of the alphabet: **Other alphabets** (e.g. Cyrillic, Greek, Chinese, Hebrew, etc.).

Not at all   A little bit   Moderately   Quite a lot   Completely

4.A11-3) I perceive different <u>colours</u> when I see, hear, or think about the **signs of sign languages**.

Not at all   A little bit   Moderately   Quite a lot   Completely

5.A11-4) I perceive different <u>colours</u> when I see, hear, or think about **words**.

Not at all   A little bit   Moderately   Quite a lot   Completely

6.A11-5) I perceive different <u>colours</u> when I see, hear, or think about **foreign words**.

Not at all   A little bit   Moderately   Quite a lot   Completely

7.A11-6) I perceive different <u>colours</u> when I see, hear, or think about **people's names or other proper nouns** as.

Not at all   A little bit   Moderately   Quite a lot   Completely

8.A11-7) I perceive different <u>colours</u> when I see, hear, or think about **shapes** (e.g. triangle, square, etc.).

Not at all   A little bit   Moderately   Quite a lot   Completely

9.A11-8) I perceive different <u>colours</u> when I see, hear, or think about **punctuation and/or other known symbols** (e.g. !, &, etc.).

Not at all   A little bit   Moderately   Quite a lot   Completely

10.A11-9) I perceive different <u>colours</u> when I see, hear, or think about **numbers**.

Not at all   A little bit   Moderately   Quite a lot   Completely

11.A11-10) I perceive different <u>colours</u> when I see, hear, or think about the **days of the week**.

Not at all   A little bit   Moderately   Quite a lot   Completely

12.A11-11) I perceive different <u>colours</u> when I see, hear, or think about the **months of the year**.

Not at all   A little bit   Moderately   Quite a lot   Completely

13.A11-12) I perceive different <u>colours</u> when I see, hear, or think about the **years and/or decades, centuries, etc**.

Not at all   A little bit   Moderately   Quite a lot   Completely

14.A11-13) I perceive different <u>colours</u> when I see, hear, or think about the **hours of the day**.

Not at all   A little bit   Moderately   Quite a lot   Completely

15.A11-14) I perceive different <u>colours</u> when I see, hear, or think about **measurement scales** (e.g. height, weight, temperatures, etc.).

Not at all   A little bit   Moderately   Quite a lot   Completely

16.A12) I perceive different <u>colours</u> when I listen to or hear music, sounds/noises, or voices. For example: *When I listen to pop music, I see pink to red tonalities*, or *When I hear my mother speaking, I see turquoise shades*.

Not at all A little bit Moderately Quite a lot Completely

If you have answered *Not at all*, skip to Q22.

17.A12-1) I perceive different colours when I listen to or hear **musical pitches**.

Not at all A little bit Moderately Quite a lot Completely

18.A12-2) I perceive different colours when I listen to or hear **musical chords**.

Not at all A little bit Moderately Quite a lot Completely

19.A12-3) I perceive different colours when I listen to or hear **musical instruments**.

Not at all A little bit Moderately Quite a lot Completely

20.A12-4) I perceive different colours when I listen to or hear **sounds/noises**.

Not at all A little bit Moderately Quite a lot Completely

21.A12-5) I perceive different colours when I listen to or hear **voices**.

Not at all A little bit Moderately Quite a lot Completely

22.A13) I perceive different <u>colours</u> in response to sensations such as emotions, people/personalities, orgasms, touch, pain, body postures, tastes, smells, or motion/movement. For example: *The colour orange comes to me for happy experiences*, or *I see dark blue flashes when I am in pain*.

Not at all A little bit Moderately Quite a lot Completely

If you have answered *Not at all*, skip to Q32.

23.A13-1) I perceive different colours in response to **emotions**.

Not at all A little bit Moderately Quite a lot Completely

24.A13-2) I perceive different colours in response to **people/personalities**.

Not at all A little bit Moderately Quite a lot Completely

25.A13-3) I perceive different colours in response to **orgasms**.

Not at all A little bit Moderately Quite a lot Completely

26.A13-4) I perceive different colours in response to **touch sensations**.

Not at all A little bit Moderately Quite a lot Completely

27.A13-5) I perceive different colours in response to **pain sensations**.

Not at all A little bit Moderately Quite a lot Completely

28.A13-6) I perceive different colours in response to **body postures**.

Not at all A little bit Moderately Quite a lot Completely

29.A13-7) I perceive different colours in response to **taste sensations**.

Not at all A little bit Moderately Quite a lot Completely

30.A13-8) I perceive different colours in response to **smell sensations**.

Not at all A little bit Moderately Quite a lot Completely

31.A13-9) I perceive different colours in response to **motion/movement**.

Not at all A little bit Moderately Quite a lot Completely

32.A21) I see different <u>visual patterns</u> such as waves, lines, or circles – with or without movement, when I hear or listen to music, sounds/noises, or voices; or in response to the experience of emotions, orgasms, people/personalities, touch, pain, body postures, taste, smell, or motion/movement. For example: *When I listen to classical music, I see a succession of wavy lines intermixed with abstract shapes that follow the melody*, or *When I experience happy emotions I see round shapes but for angry ones they are pointy*.

Not at all A little bit Moderately Quite a lot Completely

If you have answered *Not at all*, skip to Q45.

33.A21-1) I see different visual patterns when I listen to or hear **music**.

Not at all A little bit Moderately Quite a lot Completely

34.A21-2) I see different visual patterns when I listen to or hear **sounds/noises**.

Not at all A little bit Moderately Quite a lot Completely

35.A21-3) I see different visual patterns when I listen to or hear **voices**.

Not at all A little bit Moderately Quite a lot Completely

36.A21-4) I see different visual patterns in response to **emotions**.

Not at all A little bit Moderately Quite a lot Completely

37.A21-5) I see different visual patterns in response to **people/personalities**.

Not at all A little bit Moderately Quite a lot Completely

38.A21-6) I see different visual patterns in response to **orgasms**.

Not at all A little bit Moderately Quite a lot Completely

39.A21-7) I see different visual patterns in response to **touch sensations**.

Not at all A little bit Moderately Quite a lot Completely

40.A21-8) I see different visual patterns in response to **pain sensations**.

Not at all A little bit Moderately Quite a lot Completely

41.A21-9) I see different visual patterns in response to **body postures**.

Not at all A little bit Moderately Quite a lot Completely

42.A21-10) I see different visual patterns in response to **taste sensations**.

Not at all A little bit Moderately Quite a lot Completely

43.A21-11) I see different visual patterns in response to **smell sensations**.

Not at all A little bit Moderately Quite a lot Completely

44.A21-12) I see different visual patterns in response to **motion/movement**.

Not at all A little bit Moderately Quite a lot Completely

45.A22) When I hear someone speaking or when I speak or think to myself, I see the <u>words spelled out</u> as a scrolling “prompter” or “subtitles” would look like in a way.

Not at all A little bit Moderately Quite a lot Completely

46.A23) I see the alphabet, numbers, weekdays, months, years, hours, measurement scales, or musical notes arranged in <u>specific patterns or shapes</u> such as lines, circles, or blocks. For example: W*hen I think what I am going to do next summer, I see the months arranged counter-clockwise in a sort of oval shape*.

Not at all A little bit Moderately Quite a lot Completely

If you have answered *Not at all*, skip to Q55.

47.A23-1) I see the **letters of the alphabet** arranged in a specific pattern or shape such as a line, a circle, blocks, etc.

Not at all A little bit Moderately Quite a lot Completely

48.A23-2) I see the **numbers** arranged in a specific pattern or shape such as a line, a circle, blocks, etc.

Not at all A little bit Moderately Quite a lot Completely

49.A23-3) I see the **days of the week** arranged in a specific pattern or shape such as a line, a circle, blocks, etc.

Not at all A little bit Moderately Quite a lot Completely

50.A23-4) I see the **months of the year** arranged in a specific pattern or shape such as a line, a circle, blocks, etc.

Not at all A little bit Moderately Quite a lot Completely

51.A23-5) I see the **years and/or decades, centuries, etc.** arranged in a specific pattern or shape such as a line, a circle, blocks, etc.

Not at all A little bit Moderately Quite a lot Completely

52.A23-6) I see the **hours of the day** arranged in a specific pattern or shape such as a line, a circle, blocks, etc.

Not at all A little bit Moderately Quite a lot Completely

53.A23-7) I see **measurement scales** (e.g. height, weight, temperatures, etc.) arranged in specific patterns or shapes such as lines, circles, blocks, etc.

Not at all A little bit Moderately Quite a lot Completely

54.A23-8) I see the **notes of musical scales** arranged in a specific or shape such as a line, a circle, blocks, etc.

Not at all A little bit Moderately Quite a lot Completely

55.B11/12) I experience PHYSICAL <u>touch or pain sensations</u> in response to words, concepts, music, sounds, voices, emotions, people/personalities, or colours. For example: *When I listen to music, I experience* physical *touch sensations in the back of my head*, or *If I read in a book that someone is being punched in the stomach, I experience* physical *pain sensations in my own stomach*.

Not at all A little bit Moderately Quite a lot Completely

If you have answered *Not at all*, skip to Q63.

56a.B11/12-1) I experience PHYSICAL touch or pain sensations in response to **concepts** such as the letters of the alphabet, words, numbers, weekdays, shapes, symbols, etc.

Not at all A little bit Moderately Quite a lot Completely

56b.B11/12-1) If you answered anything different from *Not at all*, please specify:

Only touch sensations Only pain sensations Both touch and pain sensations

57a.B11/12-2) I experience PHYSICAL touch or pain sensations in response to **music**.

Not at all A little bit Moderately Quite a lot Completely

57b.B11/12-2) If you answered anything different from *Not at all*, please specify:

Only touch sensations Only pain sensations Both touch and pain sensations

58a.B11/12-3) I experience PHYSICAL touch or pain sensations in response to **sounds/noises**.

Not at all A little bit Moderately Quite a lot Completely

58b.B11/12-3) If you answered anything different from *Not at all*, please specify:

Only touch sensations Only pain sensations Both touch and pain sensations

59a.B11/12-4) I experience PHYSICAL touch or pain sensations in response to **voices**.

Not at all A little bit Moderately Quite a lot Completely

59b.B11/12-4) If you answered anything different from *Not at all*, please specify:

Only touch sensations Only pain sensations Both touch and pain sensations

60a.B11/12-5) I experience PHYSICAL touch or pain sensations in response to **emotions**.

Not at all A little bit Moderately Quite a lot Completely

60b.B11/12-5) If you answered anything different from *Not at all*, please specify:

Only touch sensations Only pain sensations Both touch and pain sensations

61a.B11/12-6) I experience PHYSICAL touch or pain sensations in response to **people/personalities**.

Not at all A little bit Moderately Quite a lot Completely

61b.B11/12-6) If you answered anything different from *Not at all*, please specify:

Only touch sensations Only pain sensations Both touch and pain sensations

62a.B11/12-7) I experience PHYSICAL touch or pain sensations in response to **colours**.

Not at all A little bit Moderately Quite a lot Completely

62b.B11/12-7) If you answered anything different from *Not at all*, please specify:

Only touch sensations Only pain sensations Both touch and pain sensations

63.B13) I experience ACTUAL taste sensations in response to words, concepts, music, sounds, voices, emotions, people/personalities, colours, orgasms, touch, pain, body postures, or motion/movement. For example: *When I hear or think about the word chair, I experience an* actual *tangy apple taste in my mouth*.

Not at all A little bit Moderately Quite a lot Completely

If you have answered *Not at all*, skip to Q76.

64.B13-1) I experience ACTUAL taste sensations in response to **concepts** such as the letters of the alphabet, words, numbers, weekdays, shapes, symbols, etc.

Not at all A little bit Moderately Quite a lot Completely

65.B13-2) I experience ACTUAL taste sensations in response to **music**.

Not at all A little bit Moderately Quite a lot Completely

66.B13-3) I experience ACTUAL taste sensations in response to **sounds/noises**.

Not at all A little bit Moderately Quite a lot Completely

67.B13-4) I experience ACTUAL taste sensations in response to **voices**.

Not at all A little bit Moderately Quite a lot Completely

68.B13-5) I experience ACTUAL taste sensations in response to **emotions**.

Not at all A little bit Moderately Quite a lot Completely

69.B13-6) I experience ACTUAL taste sensations in response to **people/personalities**.

Not at all A little bit Moderately Quite a lot Completely

70.B13-7) I experience ACTUAL taste sensations in response to **colours**.

Not at all A little bit Moderately Quite a lot Completely

71.B13-8) I experience ACTUAL taste sensations in response to **orgasms**.

Not at all A little bit Moderately Quite a lot Completely

72.B13-9) I experience ACTUAL taste sensations in response to **touch sensations**.

Not at all A little bit Moderately Quite a lot Completely

73.B13-10) I experience ACTUAL taste sensations in response to **pain sensations**.

Not at all A little bit Moderately Quite a lot Completely

74.B13-11) I experience ACTUAL taste sensations in response to **body postures**.

Not at all A little bit Moderately Quite a lot Completely

75.B13-12) I experience ACTUAL taste sensations in response to **motion/movement**.

Not at all A little bit Moderately Quite a lot Completely

76.B14) I experience ACTUAL smell sensations in response to words, concepts, music, sounds, voices, emotions, people/personalities, colours, orgasms, touch, pain, body postures, or motion/movement. For example: *When I am sad, I experience an* actual *mouldy scent in my nose*.

Not at all A little bit Moderately Quite a lot Completely

If you have answered *Not at all*, skip to Q89.

77.B14-1) I experience ACTUAL smell sensations in response to **concepts** such as the letters of the alphabet, words, numbers, weekdays, shapes, symbols, etc.

Not at all A little bit Moderately Quite a lot Completely

78.B14-2) I experience ACTUAL smell sensations in response to **music**.

Not at all A little bit Moderately Quite a lot Completely

79.B14-3) I experience ACTUAL smell sensations in response to **sounds/noises**.

Not at all A little bit Moderately Quite a lot Completely

80.B14-4) I experience ACTUAL smell sensations in response to **voices**.

Not at all A little bit Moderately Quite a lot Completely

81.B14-5) I experience ACTUAL smell sensations in response to **emotions**.

Not at all A little bit Moderately Quite a lot Completely

82.B14-6) I experience ACTUAL smell sensations in response to **people/personalities**.

Not at all A little bit Moderately Quite a lot Completely

83.B14-7) I experience ACTUAL smell sensations in response to **colours**.

Not at all A little bit Moderately Quite a lot Completely

84.B14-8) I experience ACTUAL smell sensations in response to **orgasms**.

Not at all A little bit Moderately Quite a lot Completely

85.B14-9) I experience ACTUAL smell sensations in response to **touch sensations**.

Not at all A little bit Moderately Quite a lot Completely

86.B14-10) I experience ACTUAL smell sensations in response to **pain sensations**.

Not at all A little bit Moderately Quite a lot Completely

87.B14-11) I experience ACTUAL smell sensations in response to **body postures**.

Not at all   A little bit   Moderately   Quite a lot   Completely

88.B14-12) I experience ACTUAL smell sensations in response to **motion/movement**.

Not at all   A little bit   Moderately   Quite a lot   Completely

89.B15) I perceive ACTUAL sounds or music in response to words, concepts, emotions, people/personalities, colours, orgasms, touch, pain, body postures, taste, smell, or motion/movement. For example: *When I see moving dots in a screen, I hear* actual *beeps or taps*.

Not at all   A little bit   Moderately   Quite a lot   Completely

If you have answered *Not at all*, skip to Q101.

90.B15-1) I perceive ACTUAL sounds or music in response to **concepts** such as the letters of the alphabet, words, numbers, weekdays, shapes, symbols, etc.

Not at all   A little bit   Moderately   Quite a lot   Completely

91.B15-2) I perceive ACTUAL sounds or music in response to **emotions**.

Not at all   A little bit   Moderately   Quite a lot   Completely

92.B15-3) I perceive ACTUAL sounds or music in response to **people/personalities**.

Not at all   A little bit   Moderately   Quite a lot   Completely

93.B15-4) I perceive ACTUAL sounds or music in response to **colours**.

Not at all   A little bit   Moderately   Quite a lot   Completely

94.B15-5) I perceive ACTUAL sounds or music in response to **orgasms**.

Not at all   A little bit   Moderately   Quite a lot   Completely

95.B15-6) I perceive ACTUAL sounds or music in response to **touch sensations**.

Not at all   A little bit   Moderately   Quite a lot   Completely

96.B15-7) I perceive ACTUAL sounds or music in response to **pain sensations**.

Not at all   A little bit   Moderately   Quite a lot   Completely

97.B15-8) I perceive ACTUAL sounds or music in response to **body postures**.

Not at all   A little bit   Moderately   Quite a lot   Completely

98.B15-9) I perceive ACTUAL sounds or music in response to **taste sensations**.

Not at all   A little bit   Moderately   Quite a lot   Completely

99.B15-10) I perceive ACTUAL sounds or music in response to **smell sensations**.

Not at all   A little bit   Moderately   Quite a lot   Completely

100.B15-11) I perceive ACTUAL sounds or music in response to **motion/movement**.

Not at all   A little bit   Moderately   Quite a lot   Completely

101.B2) When I see something/someone else being touched or in (physical) pain, or when I see someone smelling or tasting something; I experience the same PHYSICAL sensation in my own body. For example: *When I see someone bumping his/her knee on a table, I experience a* physical *sensation of pain in my own knee*.

Not at all   A little bit   Moderately   Quite a lot   Completely

If you have answered *Not at all*, skip to Q106.

102.B2-1) When I see something or someone else being touched, I experience the same PHYSICAL **touch sensations** in my own body.

Not at all A little bit Moderately Quite a lot Completely

103.B2-2) When I see someone else in (physical) pain, I experience the same PHYSICAL **pain sensations** in my own body.

Not at all A little bit Moderately Quite a lot Completely

104.B2-3) When I see someone else tasting something, I experience the same ACTUAL **taste sensations** in my mouth.

Not at all A little bit Moderately Quite a lot Completely

105.B2-4) When I see someone else smelling something, I experience the same ACTUAL **smell sensations** in my nose.

Not at all A little bit Moderately Quite a lot Completely

106.C) I think about concepts such as letters, numbers, time sequences (.e.g. weekdays, months, etc.), or objects as having specific personalities or genders. For example: *To me, number 5 is not trustworthy*, or *To me, the letter L is a girly girl*. Note 1: Regarding gender attributions, they must be unrelated to the linguistic genders languages such as Spanish or French inherently present - e.g. in Spanish, 'a car' is masculine ("el coche"), whereas in French is feminine ("la voiture"). Note 2: Regarding personality or gender attributions to objects, not taking into account those objects to which you feel a strong personal attachment.

Not at all A little bit Moderately Quite a lot Completely

If you have answered *Not at all*, skip to Section B.

107.C-1) I think about the **letters of the alphabet** as having distinct personalities.

Not at all A little bit Moderately Quite a lot Completely

108.C-2) I think about the **letters of the alphabet** as having distinct genders.

Not at all A little bit Moderately Quite a lot Completely

109.C-3) I think about **numbers** as having distinct personalities.

Not at all A little bit Moderately Quite a lot Completely

110.C-4) I think about **numbers** as having distinct genders.

Not at all A little bit Moderately Quite a lot Completely

111.C-5) I think about **time sequences** (e.g. weekdays, months, years, hours, etc.) as having distinct personalities.

Not at all A little bit Moderately Quite a lot Completely

112.C-6) I think about **time sequences** (e.g. weekdays, months, years, hours, etc.) as having distinct genders.

Not at all A little bit Moderately Quite a lot Completely

113.C-7) I think about **objects** as having distinct personalities (besides those objects to which you feel a strong personal attachment).

Not at all A little bit Moderately Quite a lot Completely

114.C-8) I think about **objects** as having distinct genders (besides those objects to which you feel a strong personal attachment).

Not at all A little bit Moderately Quite a lot Completely

115) Do you have other similar experiences you would like to mention? If so, please specify and rate each of them.

E.g. letters-colours (i.e. letters cause the experience of colours)	Not at all	A little bit	Moderately	Quite a lot	Completely
	Not at all	A little bit	Moderately	Quite a lot	Completely
	Not at all	A little bit	Moderately	Quite a lot	Completely
	Not at all	A little bit	Moderately	Quite a lot	Completely
	Not at all	A little bit	Moderately	Quite a lot	Completely
	Not at all	A little bit	Moderately	Quite a lot	Completely
	Not at all	A little bit	Moderately	Quite a lot	Completely

Appendix E: Studies 1 and 3 Exploratory Analyses

Additional synaesthetes (Study 3)

In the following analyses, we examined whether the behavioural performance of alternative groups of synaesthetes (i.e. both- and other-synaesthetes) compared to the synaesthetes of interest examined in the main analyses of Study 3 (see section 2.4.3). To do these exploratory analyses, we conducted mixed analyses of variance (ANOVAs) on reaction times (RT), separately for each task. Given the differences observed in the main analyses between colour- and sequence-synaesthetes, we decided to keep the subsamples separated. For each ANOVA; ‘Congruency’ (congruent, incongruent) was the within-subjects factor and ‘Group’ (non-synaesthetes, colour-synaesthetes, sequence-synaesthetes, both/other-synaesthetes) the between-subjects factor. Further independent *t*-tests on the congruency effects or CE (i.e. differences between incongruent and congruent trials) were carried out as appropriate following significant interactions. We ommitted the comparisons already examined in the main analyses (i.e. non- vs. colour-synaesthetes, non- vs. sequence-synaesthetes, and colour- vs. sequence-synaesthetes) and thus defined the following post-hoc target comparisons: both/other-synaesthetes vs. non-synaesthetes, vs. colour-synaesthetes, and vs. sequence-synaesthetes.

The specifications regarding the inclusion and exclusion of participants and trials in RT analyses followed the same criteria detailed in Study 1 (see section 2.2.2.3). For all analyses, in case of violation of the assumption of normal distribution of the dependent variables (as assessed by Shapiro-Wilk) we used Mann-Whitney U tests (we did not adopt any further non-parametric approaches as the control analyses run in the main analyses to address this issue did not reveal differences in significance). In addition, Bonferroni corrections for multiple comparisons were applied adjusting the alpha threshold accordingly when needed**.** Lastly, we only analysed RT data as any group effects observed in the main analyses emerged in terms of differences in milliseconds and not error rates. The analyses were conducted in Jamovi 0.9 (Jamovi Project, 2018) and SPSS 24 (IBM Corporation, 2016).

Both-synaesthetes analyses

In the Visuo-Tactile Concurrent-Unrelated (VTCU) task, the main effect for the factor 'Group' was not significant ($F(3, 74) = .82$, $p = .49$), but there was a significant interaction between 'Congruency' and 'Group' ($F(3, 74) = 2.78$, $p = .047$, $\eta_p^2 = .10$). However, follow-up independent-samples tests for the target comparisons indicated no differences between both-synaesthetes (M = 17.2, SD = 20.3) and the other groups: non-synaesthetes (M = 29, SD = 25.6; $U(35) = 106$, $p = .23$), colour-synaesthetes (M = 20, SD = 19.2; $t(30) = .38$, $p = .71$), and sequence-synaesthetes (M = 11.5, SD = 14.6; $t(29) = .90$, $p = .38$) (Bonferroni-adjusted $\alpha = .017$). There were not main effects of 'Group' or interactions involving this factor for any of the rest of the tasks (all $F(3, 74) < 1.18$, all $p > .32$) (Fig. 9 in section 2.4.4).

Other-synaesthetes analyses

The main effect of 'Group' was not significant in the VTCU task ($F(3, 85) = .63$, $p = .60$), but there was a significant interaction between 'Congruency' and 'Group' ($F(3, 85) = 3.60$, $p = .017$, $\eta_p^2 = .11$). Follow-up independent-samples tests for the target comparisons with Bonferroni correction for multiple comparisons ($\alpha = .017$) showed that other-synaesthetes presented significantly smaller CE than non-synaesthetes in this task (M = 12.8, SD = 19.6 and M = 29, SD = 25.6, respectively; $U(46) = 152$, $p = .005$, $d = .70$). However, no differences were found between other-synaesthetes and colour- and sequence-synaesthetes (colour-synaesthetes: M = 20, SD = 19.2; $t(41) = 1.20$, $p = .24$ and sequence-synaesthetes: M = 11.5, SD = 14.6; $t(40) = .24$, $p = .81$). There were not main effects of 'Group' or interactions concerning this factor for any of the rest of the tasks (all $F(3, 85) < 1.53$, all $p > .21$) (Fig. 10 in section 2.4.4).

Synaesthetic strength (Studies 1 and 3)

We conducted exploratory Pearson correlations (or Spearman rank correlations for ordinal variables; e.g. Salkind, 2010; Schober et al., 2018) to examine whether several synaesthesia strength measures were related to the congruency effects or CE (i.e. differences between incongruent and congruent trials) of those tasks for which group differences were observed (i.e. Study 1: CCT, Study 3: VTCU). In particular three synaesthetic strength measures were assessed: number of synaesthesia types reported, overall degree of synaesthetic experience reported, and synaesthetic consistency scores. Number of synaesthesia types and overall degree of synaesthetic experience were extracted from the ESSA (Edinburgh Synaesthesia Screening Assessment) interview responses. We chose the ESSA Extended Highest score to measure the overall degree of synaesthetic experience because this was the best performing scoring in our questionnaire study (see Chapter IV). See section 2.5 for further details on the strength measures defined and Table E1 for a summary of the analyses performed and their sources.

As we were interested in assessing whether synaesthesia strength had any moderating effect in behavioural performance, these analyses were only conducted on the synaesthetes subsamples since, by definition, non-synaesthetes do not experience any degree of synaesthetic experience. This distinction is intrinsic of the synaesthetic strength measures assessed. That is, although these measures provide scores on a continuum, the differentiation between synaesthetes and non-synaesthetes is categorical as it is defined on thresholds [of those continuums]. Thus, if we included non-synaesthetes in the analyses, any relationships found between synaesthesia strength and CE could not be distinguished from this qualitative difference. Following the same logic, in Study 3 -colour and sequence consistency scores were only analysed with the respective colour- and sequence-synaesthetes subsamples (as sequence-synaesthetes present ‘non-synaesthetic’ -colour scores and vice-versa for colour-synaesthetes in relation to sequence scores). Given the differences observed in the main analyses between these two groups (see section 2.4.3), we decided to keep the subsamples separated for all the analyses. Lastly, it should be noted that the overall degree of synaesthetic experience could only be evaluated in Study 3 as this measure was obtained from the new version of the screening assessment (revised ESSA) and was not available in the old version used in Study 1 (pilot ESSA). In addition, we did not perform Study 1 -colour consistency analyses because we considered that subsample was too small (*N* = 9) to produce any meaningful results.

Bonferroni corrections were applied for multiple comparisons (adjusted $\alpha = .017$). We focused on reaction times (RT) as any group effects observed in the main analyses emerged in terms of differences in milliseconds and not error rates. In addition, although there were some problems with the normality of the data in the different studies, all the control non-parametric analyses run did not provide alternative conclusions and ratified the main analyses. Therefore, the use of non-parametric approaches was not deemed necessary for the present analyses. The analyses were conducted in Jamovi 0.9 (Jamovi Project, 2018).

Table E1.
Summary of the different synaesthetic strength measures assessed in Studies 1 and 3 and their sources.

Synaesthetic strength measure	Study 1	Study 3
Number of synaesthesia types (self-report)	Pilot ESSA	Revised ESSA (separate analyses for colour- and sequence-synaesthetes subsamples)
Overall degree of synaesthetic experience (self-report)	–	Revised ESSA – Extended Highest score (separate analyses for -colour and sequence synaesthetes subsamples)
-Colour consistency scores (consistency test)	–	Synesthesia Battery/Multisense Consistency Test (-colour synaesthetes subsample only)*
Sequence consistency scores (consistency test)	–	Sussex's Sequence-Spatial Synaesthesia Diagnostic Test – consistency test (sequence synaesthetes subsample only)

Note: ESSA = Edinburgh Synaesthesia Screening Assessment.
* Study 3 scores of participants who completed the Multisense Consistency Test were transformed to Synesthesia Battery scores for homogenisation purposes.

Study 1

The analysis revealed that the number of synaesthesia types self-reported by synaesthetes and the CE of the CCT task was not related (r = -.010, p = .97; Bonferroni-adjusted α = .017).

Study 3

No relationships were found between the number of synaesthesias, the overall degree of synaesthetic experience, or synaesthetic consistency scores and the CE size of Study's 3 VTCU task of the colour- or sequence-synaesthetes subsamples (all r > .35, all p < .13; Bonferroni-adjusted α = .017; see Table E2 for details).

Table E2.
Correlations between the congruency effects and the synaesthetic strength measures analysed in Studies 1 and 3, by group.

Synaesthetic strength measure	*N*		CCT	VTCU
Study 1				
P-ESSA nº of synaesthesia types	16	Pearson's *r* / *p*-value	-.010 / .97	-
Study 3				
Colour-synaesthetes				
R-ESSA nº of synaesthesia types	21	Pearson's *r* / *p*-value	-	.15 / .51
R-ESSA degree of synaesthetic experience	21	Spearman's ρ / *p*-value	-	.17 /.45
-Colour SB/MCT consistency score*	21	Pearson's *r* / *p*-value	-	-.25 / .28
Sequence-synaesthetes				
R-ESSA nº of synaesthesia types	20	Pearson's *r* / *p*-value	-	.066 / .78
R-ESSA degree of synaesthetic experience	20	Spearman's ρ / *p*-value	-	.35 / .13
Sequence SDT consistency score	20	Pearson's *r* / *p*-value	-	-.010 / .97

Note: Congruency effects = Reaction time differences in milliseconds between incongruent and congruent trials; *N* = Sample size; CCT = Cross-modal Congruency Task; VTCU = Visuo-Tactile Concurrent-Unrelated task; SB = Syncesthesia Battery; MCT = Multisense Consistency Test; SDT = Sussex's Sequence-Spatial Synaesthesia Diagnostic Test; ESSA = Edinburgh Synaesthesia Screening Assessment (P = Pilot; R = Revised).
* Scores of participants who completed the MCT were transformed to SB scores for homogenisation purposes.

Appendix F: Study 4 Exploratory Analyses

In the following analyses, we examined whether certain personality trait scores of alternative groups of synaesthetes (i.e. both- and other-synaesthetes) compared to the scores of the synaesthetes of interest examined in the main analyses of Study 4 (see section 3.3). To do these exploratory analyses, the mean scores on each personality trait assessed were submitted to Analyses of Variance (ANOVAs) with 'Group' (non-synaesthetes, colour-synaesthetes, sequence-synaesthetes, both/other-synaesthetes) as the fixed factor, separately for each sample (there were no both-synaesthetes in Sample B, therefore, these analyses were only performed for Sample A). Further post-hoc independent samples *t*-tests were carried out as appropriate following significant main effects. We ommitted the comparisons already examined in the main analyses (i.e. non- vs. colour-synaesthetes, non- vs. sequence-synaesthetes, and colour- vs. sequence-syanesthetes) and thus defined the following post-hoc target comparisons: both/other-synaesthetes vs. non-synaesthetes, vs. colour-synaesthetes, and vs. sequence-synaesthetes. In case of violation of the assumption of normal distribution of the dependent variables (as assessed by Shapiro-Wilk), we used alternative, non-parametric Mann-Whitney U tests (we did not adopt any further non-parametric approaches as the control analyses run in the main analyses to address this issue did not reveal differences in significance). Bonferroni corrections for multiple comparisons were applied adjusting the alpha threshold accordingly when needed. The analyses were conducted in Jamovi 0.9 (Jamovi Project, 2018) and SPSS 24 (IBM Corporation, 2016).

Both-synaesthetes

Sample A

The analyses revealed significant group differences for the BFI (Big Five Inventory) Openness to Experience subscale ($F(3, 113) = 6.72$, $p < .001$, $\eta_p^2 = .15$). Post-hoc comparisons with Bonferroni correction ($\alpha = .017$) indicated that both-synaesthetes (M = 3.84, SD = .60) experienced significantly higher rates than non-synaesthetes (M = 3.37 , SD = .47; $t(56)$ =

3.36, p = .001, d = .89; Fig. 13A in section 3.4). No differences were observed between both-synaesthetes and colour-synaesthetes (M = 3.59, SD = .47; t(66) = 1.92, p = .060) or sequence-synaesthetes (M = 3.96, SD = .54; t(43) = .69, p = .50). We also explored the four other subscales of the BFI with the alpha adjusted for multiple comparisons (α = .013), but no group differences were observed for any of the traits (all F(3, 113) < 3.18, all p > .027). Regarding the IRI (Interpersonal Reactivity Index) Fantasising subscale, group differences only approached significances (F(3, 113) = 2.45, p = .067). No differences were found for the rest of the IRI traits (all F(3, 113) < 1.61, all p > .19; Bonferroni-adjusted α = .017). Lastly, the analyses revealed significant group differences for the Unusual Experiences subscale of the O-LIFE (Oxford-Liverpool Inventory of Feelings and Experiences) (F(3, 113) = 8.26, p < .001, η_p^2 = .18; Bonferroni-adjusted α = .025). Post-hoc comparisons determined that both-synaesthetes (M = .46, SD = .23) presented significantly higher rates than non-synaesthetes (M = .19, SD = .17; t(56) = 5.10, p < .001, d = 1.34; Fig. 13B in section 3.4). The same pattern was observed with respect to colour-synaesthetes (M = .34, SD = .23), but the comparison did not survive corrections (t(66) = 2.04, p = .046). Both- and sequence-synaesthetes (M = .37, SD = .19) did not differ in their scores (t(43) = 1.31, p = .20). There were no group differences either for the O-LIFE Cognitive Disorganisation subscale (F(3, 113) = .79, p = .51).

Other-synaesthetes

Sample A

The analyses revealed significant group differences for the BFI Openness to Experience subscale (F(3, 125) = 5.65, p = .001, η_p^2 = .12). Post-hoc comparisons with Bonferroni correction (α = .017) indicated that other-synaesthetes (M = 3.70, SD = .55) experienced significantly higher rates than non-synaesthetes (M = 3.37 , SD = .47; t(68) = 2.65, p = .010, d = .64; Fig. 14A in section 3.4). No differences were observed between other-synaesthetes and colour-synaesthetes (M = 3.59, SD = .47; t(78) = .92, p = .36) or sequence-synaesthetes (M = 3.96, SD = .54; t(55) = 1.70, p = .095). We also explored the four other subscales of the

BFI with the alpha adjusted for multiple comparisons (α = .013), but no group differences were observed for any of the traits (all $F(3, 125) < 2.30$, all $p > .081$). Group differences only approached significances for the IRI Fantasising subscale ($F(3, 125) = 2.32$, $p = .079$), and no differences were found for the rest of the IRI traits (all $F(3, 125) < 1.56$, all $p > .20$; Bonferroni-adjusted α = .017). Regarding the O-LIFE traits, significant group differences were found for the Unusual Experiences subscale ($F(3, 125) = 4.75$, $p = .004$, $\eta_p^2 = .10$; Bonferroni-adjusted α = .025). Post-hoc comparisons determined that other-synaesthetes (M = .37, SD = .27) presented significantly higher rates than non-synaesthetes (M = .19, SD = .17; $U(68) = 348$, $p = .002$, $d = .81$; Fig. 14B in section 3.4). No differences were observed between other-synaesthetes and colour-synaesthetes (M = .34, SD = .23; $t(78) = .45$, $p = .66$) and sequence-synaesthetes (M = .37, SD = .19; $t(55) = .09$, $p = .93$). There were no group differences either for the O-LIFE Cognitive Disorganisation subscale ($F(3, 125) = 1.67$, $p = .18$).

Sample B

Group differences in the BFI Openness to Experiences subscale only approach significance ($F(3, 271) = 2.46$, $p = .063$). There were neither differences for the other subscales of the BFI (all $F(3, 271) < 1.48$, all $p > .22$; Bonferroni-adjusted α = .013). In relation to the IRI traits, no differences were observed for Fantasising ($F(3, 275) = 1.12$, $p = .34$) or the rest of the traits (all $F(3, 271) < .61$, all $p > .61$; Bonferroni-adjusted α = .017). Lastly, the analyses revealed significant group differences for the O-LIFE Unusual Experiences subscale ($F(3, 271) = 10.3$, $p < .001$, $\eta_p^2 = .10$; Bonferroni-adjusted α = .025). Post-hoc comparisons determined that other-synaesthetes (M = .46, SD = .20) presented significantly higher rates than non-synaesthetes (M = .30, SD = .19; $t(211) = 5.61$, $p < .001$, $d = .80$) and colour-synaesthetes (M = .33, SD = .23; $t(129) = 3.51$, $p < .001$, $d = .63$) (Fig. 15 in section 3.4). Other- and sequence-synaesthetes (M = .35, SD = .18) did not differ in their scores ($t(87) = 1.69$, $p = .094$). Group differences only approach significance for the O-LIFE Cognitive Disorganisation subscale ($F(3, 271) = 2.25$, $p = .083$).